

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2002, 09:37:55 ; Search time 11.82 Seconds
(without alignments)
360.334 Million cell updates/sec

Title: US-10-019-455A-24

Perfect score: 590
Sequence: 1 HGIPMDRLASKKICADDECV.....RVYQETKVEPTTIDIDFFCE 110

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	590	100.0	128	1	OTOR_HUMAN
2	547	92.7	128	1	OTOR_MOUSE
3	477.5	80.9	132	1	OTOR_CHICK
4	390	66.1	133	1	OTOR_RANCA
5	253.5	43.0	131	1	MIA_HUMAN
6	252.5	42.8	130	1	MIA_BOVIN
7	250.5	42.5	130	1	MIA_RAT
8	235.5	39.9	130	1	MIA_MOUSE
9	88.5	15.0	847	1	VAV3_MOUSE
10	86.5	14.7	845	1	VAV3_HUMAN
11	86.5	14.7	847	1	VAV3_HUMAN
12	85.5	14.5	843	1	VAV_MOUSE
13	85.5	14.5	845	1	VAV_MOUSE
14	82.5	14.0	868	1	VAV2_MOUSE
15	81.5	13.8	878	1	VAV2_HUMAN
16	78.5	13.3	1589	1	CC25_YEAST
17	72	12.2	520	1	ITSN_HUMAN
18	71.5	12.1	541	1	ASNH_METJA
19	70.5	11.9	905	1	ZO3_MOUSE
20	69	11.7	392	1	RURE_ACICA
21	69	11.7	444	1	PUR2_METJA
22	69	11.7	670	1	YB2_SCHPO
23	68.5	11.6	505	1	SRK1_SPOLA
24	68.5	11.6	506	1	SRK4_SPOLA
25	68	11.5	162	1	BAR2_CHIPA
26	68	11.5	259	1	YA20_METJA
27	68	11.5	534	1	VL2_HPV37
28	67	11.4	383	1	PGL3_ASPNG
29	67	11.4	509	1	SYE_VTBCH
30	67	11.4	518	1	VL2_HPV08
31	67	11.4	524	1	VL2_HPV17
32	66	11.2	326	1	BAR1_CHIPA
33	66	11.2	518	1	VL2_HPV36

ALIGNMENTS

RESULT	1	OTOR_HUMAN	STANDARD;	PRT;	128 AA.
ID	OTOR_HUMAN	STANDARD;	PRT;	128 AA.	
AC	Q9NRC9;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	01-MAR-2002 (Rel. 41, Last annotation update)				
DE	Otoraplin precursor (Fibrocyte-derived protein) (Melanoma inhibitory activity like protein).				
DE	activity like protein).				
GN	OTOR OR FDP OR MIAL.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Cochlea;				
RX	MEDLINE=20334619; PubMed=10873378;				
RA	Robertson N.G., Heller S., Lin J.S., Resendes B.L., Weremowicz S., Denis C.S., Bell A.M., Hudspeth A.J., Morton C.C.;				
RT	"A novel conserved cochlear gene, OTOR: identification, expression analysis, and chromosomal mapping."				
RL	Genomics 66:242-248(2000).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20568254; PubMed=10998416;				
RA	Cohen-Salmon M., Frenz D., Liu W., Verpy E., Voegelings S., Petit C.;				
RT	"Fdp, a new fibrocyte-derived protein related to MIA/CD-RAP, has an in vitro effect on the early differentiation of the inner ear mesenchyme."				
RL	J. Biol. Chem. 275:40036-40041(2000).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Brain, and Cochlea;				
RX	MEDLINE=21100875; PubMed=11161796;				
RA	Rendtorff N.D., Frodin M., Attie-Bitach T., Vekemans M., Tommerup N.;				
RT	"Identification and characterization of an inner ear-expressed human melanoma inhibitory activity (MIA)-like gene (MIAL) with a frequent polymorphism that abolishes translation."				
RL	Genomics 71:40-52(2001).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L., Balley J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J., Buck D., Burrill W., Butler A.P., Carder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Cobley V.E., Collier R.E., Connor R., Corby N.R., Coulson A., Coville G.J., Deadman R., Dhani P., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Lehtvaslaih M.H., Levensha M., Lloyd C., Lloyd D.M., Lovell J.D.,				

O62683 canis famil
P36568 serratia ma
Q01886 cochliobolu
P44499 haenophilus
P22425 human papil
O95049 homo sapien
P26836 clostridium
P16126 leishmania
P18791 streptococc
P34258 caenorhabdl
P19582 bacillus su
Q45755 bacillus th


```
Proc. Natl. Acad. Sci. U.S.A. 98:5515-5520(2001).
-!- FUNCTION: ELICITS GROWTH INHIBITION ON MELANOMA CELLS IN VITRO
AS WELL AS SOME OTHER NEUROECODERMAL TUMORS, INCLUDING GLIOMAS.
-!- SUBCELLULAR LOCATION: Secreted.
-!- TISSUE SPECIFICITY: ALL MALIGNANT MELANOMA CELL LINES TESTED AND
INFREQUENTLY IN GLIOMA CELL LINES.
-!- PTM: MAY POSSESS TWO INTRAMOLECULAR DISULFIDE BONDS.
-!- SIMILARITY: BELONGS TO THE MIA / OTOR FAMILY.
-!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; X75450; CAA53203.1; -
EMBL; X84707; CAA59195.1; -
EMBL; BC005910; AAH05910.1; -
PDB; 1IIJ; 16-MAY-01.
MIM; 601340; -
InterPro; IPR001452; SH3.
Pfam; PF00018; SH3; 1.
SMART; SM00326; SH3; 1.
PROSITE; PS50002; SH3; 1.
Growth factor; Signal; SH3 domain; 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 131
FT DOMAIN 43 113
FT DISULFID 36 41
FT DISULFID 59 130
FT SEQUENCE 131 AA; 14509 MW; 4D3BB30BD6008BDC CRC64;
SQ
Query Match 43.0%; Score 253.5; DB 1; Length 131;
Best Local Similarity 45.4%; Pred. No. 1.1e-19;
Matches 49; Conservative 21; Mismatches 33; Indels 5; Gaps 3;
QY 5 MDRLASKKLCADDECVYTTISLASAQEDYNAPDCRFNVKKGQIYYVSKLVKENGAGE-F 63
DB 27 MPLADRLKLCADQECSPISVAVALQDMAPDCRFITIHGQVYVYFSKL---KGRGLF 83
QY 64 WAGSVYGDGDEMGV-VGYFPRNLVKEQRYQVQATKEVPTTIDDFCE 110
DB 84 WGSVVGDDYGDLAARLGIFPSSIVREDQTLKPGKVDVTKDWDFYCQ 131
RESULT 6
MIA_BOVIN STANDARD; PRT; 130 AA.
ID MIA_BOVIN
AC Q28038;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Melanoma derived growth regulatory protein precursor (Melanoma
DE inhibitory activity) (Cartilage-derived retinoic acid-sensitive
DE protein) (CD-RAP).
GN MIA OR CDRAPI.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96216414; PubMed=8621736;
RA Dietz U.H., Sandell L.J.;
RT "Cloning of a retinoic acid-sensitive mRNA expressed in cartilage and
RL during chondrogenesis.";
RL J. Biol. Chem. 271:3311-3316(1996).
```

```
CC -!- FUNCTION: MAY FUNCTION DURING CARTILAGE DEVELOPMENT AND
CC MAINTENANCE.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: CARTILAGE PRIMORDIA AND CARTILAGE.
CC -!- INDUCTION: REPPRESSED BY RETINOIC ACID.
CC -!- PTM: MAY POSSESS TWO INTRAMOLECULAR DISULFIDE BONDS.
CC -!- SIMILARITY: BELONGS TO THE MIA / OTOR FAMILY.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; U51437; AAC48523.1; -
CC InterPro; IPR001452; SH3.
CC Pfam; PF00018; SH3; 1.
CC SMART; SM00326; SH3; 1.
CC PROSITE; PS50002; SH3; 1.
CC Growth factor; Signal; SH3 domain.
CC SIGNAL 1 22
CC CHAIN 23 130
CC DOMAIN 42 112
CC DISULFID 35 40
CC DISULFID 58 129
CC SEQUENCE 130 AA; 14353 MW; 95D153161C78E02A CRC64;
SQ
Query Match 42.8%; Score 252.5; DB 1; Length 130;
Best Local Similarity 45.0%; Pred. No. 1.4e-19;
Matches 50; Conservative 22; Mismatches 28; Indels 11; Gaps 4;
QY 5 MDRLASKKLCADDECVYTTISLASAQEDYNAPDCRFNVKKGQIYYVSKLVKENGAGE-F 63
DB 26 MPLADRLKLCADQECSPISVAVALQDMAPDCRFITIHGQVYVYFSKL---KGRGLF 82
QY 64 WAGSVYGDGDEMGV-VGYFPRNLVKEQRYQVQATKEVPTTIDDFCE 110
DB 83 WGSVVGDDYGDGAARL---GYFPSSIVREDQTLKPAKTDVTKDWDFYCQ 130
RESULT 7
MIA_RAT STANDARD; PRT; 130 AA.
ID MIA_RAT
AC Q62946; P97591;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Melanoma derived growth regulatory protein precursor (Melanoma
DE inhibitory activity) (Cartilage-derived retinoic acid-sensitive
DE protein) (CD-RAP).
GN MIA OR CDRAPI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Mammary gland;
RA Lu J.X.;
RT "Gene expression changes associated with chemically-induced rat
RT mammary carcinogenesis.";
RL Submitted (AUG-1996) to the EMBL/Genbank/DBSJ databases.
RN [2]
RP SEQUENCE OF 34-124 FROM N.A.
RC TISSUE=Cartilage;
RX MEDLINE=96216414; PubMed=8621736;
RA Dietz U.H., Sandell L.J.;
RT "Cloning of a retinoic acid-sensitive mRNA expressed in cartilage and
```



```
RT during chondrogenesis."
RL J. Biol. Chem. 271:3311-3316(1996).
CC -!- FUNCTION: MAY FUNCTION DURING CARTILAGE DEVELOPMENT AND
CC MAINTENANCE.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: CARTILAGE PRIMORDIA AND CARTILAGE.
CC -!- INDUCTION: REPPRESSED BY RETINOIC ACID.
CC -!- PTM: MAY POSSESS TWO INTRAMOLECULAR DISULFIDE BONDS.
CC -!- SIMILARITY: BELONGS TO THE MIA / OTOR FAMILY.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; U51438; AAC52481.1; -
EMBL; U67884; AAB40659.1; -
InterPro;|IPR001452; SH3.
Pfam; PF00018; SH3; 1.
SMART; SM00326; SH3; 1.
PROSITE; PS50002; SH3; FALSE_NEG.
Growth factor; Signal; SH3 domain.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 130 MELANOMA DERIVED GROWTH REGULATORY
FT DOMAIN 42 112 SH3.
FT DISULFID 35 40 BY SIMILARITY.
FT DISULFID 58 129 BY SIMILARITY.
FT CONFLICT 46 47 MA -> VT (IN REF. 2).
FT SEQUENCE 130 AA; 14536 MW; 5F99149AECF74501 CRC64;
SQ
Query Match 42.5%; Score 250.5; DB 1; Length 130;
Best Local Similarity 44.4%; Pred. No. 2.2e-19;
Matches 48; Conservative 23; Mismatches 32; Indels 5; Gaps 3;
QY 5 MDRLASKKLCADDECYVTTISLASAQEDYNAPDCRFINVKKGQIYVYSKLVKENGAGE-F 63
Db 26 MPKLDWKLCADDECSPHSMAVALQDYVAPDCRFITIRGVQVYVFSKL---KGRGRLF 82
QY 64 WAGSVYGDQDEMGV-VGFPFRNLVKEQRYQEAETKVEPTTIDIDFCE 110
Db 83 WGSVQGDYGYDLAHLGFPSPSIVREDTLTKPGKIDMKTDWDFYQ 130
ULT 8
MIA_MOUSE
ID MIA_MOUSE STANDARD; PRT; 130 AA.
AC Q61865; P97495; O09086;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Melanoma derived growth regulatory protein precursor (Melanoma
DE inhibitory activity) (Cartilage-derived retinoic acid-sensitive
DE protein) (CD-RAP).
GN MIA OR CDRAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C;
RX MEDLINE=95007612; PubMed=7923218;
RA Blesch A., Bosserhoff A.-K., Apfel R., Behl C., Hessdoerfer B.,
RA Schmitt A., Jachmnick P., Lottspeich F., Buettner R., Bogdahn U.;
RT "Cloning of a novel malignant melanoma-derived growth-regulatory
RT protein, MIA."
RL Cancer Res. 54:5695-5701(1994).
```

```
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=97251341; PubMed=9097023;
RX Bosserhoff A.K., Kondo S., Moser M., Dietz U., Copeland N.G.,
RA Gilbert D.J., Jenkins N.A., Buettner R., Sandell L.J.;
RA "Mouse CD-RAP/MIA gene: structure, chromosomal localization, and
RT expression in cartilage and chondrosarcoma.";
RL Dev. Dyn. 208:516-525(1997).
CC -!- FUNCTION: ELICITS GROWTH INHIBITION ON MELANOMA CELLS IN VITRO AS
CC WELL AS SOME OTHER NEUROECTODERMAL TUMORS, INCLUDING GLIOMAS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: ALL MALIGNANT MELANOMA CELL LINES TESTED AND
CC INFREQUENTLY IN GLIOMA CELL LINES.
CC -!- PTM: MAY POSSESS TWO INTRAMOLECULAR DISULFIDE BONDS.
CC -!- SIMILARITY: BELONGS TO THE MIA / OTOR FAMILY.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; X94322; CAA63983.1; -
EMBL; U85612; AAB42082.1; -
DR EMBL; X97965; CAA66608.1; -
DR MGD; MGI:109615; Cdrap.
DR InterPro;|IPR001452; SH3.
DR Pfam; PF00018; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50002; SH3; FALSE_NEG.
DR Growth factor; Signal; SH3 domain.
KW SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 130 MELANOMA DERIVED GROWTH REGULATORY
FT DOMAIN 42 112 SH3.
FT DISULFID 35 40 BY SIMILARITY.
FT DISULFID 58 129 BY SIMILARITY.
FT CONFLICT 112 113 TL -> NS (IN REF. 1).
FT SEQUENCE 130 AA; 14593 MW; 16C957459C5BB5F9 CRC64;
SQ
Query Match 39.9%; Score 235.5; DB 1; Length 130;
Best Local Similarity 43.5%; Pred. No. 8.1e-18;
Matches 47; Conservative 21; Mismatches 35; Indels 5; Gaps 3;
QY 5 MDRLASKKLCADDECYVTTISLASAQEDYNAPDCRFINVKKGQIYVYSKLVKENGAGE-F 63
Db 26 MPKLDWKLCADDECSPHSMAVALQDYVAPDCRFITIRGVQVYVFSKL---KGRGRLF 82
QY 64 WAGSVYGDQDEMGV-VGFPFRNLVKEQRYQEAETKVEPTTIDIDFCE 110
Db 83 WGSVQGDYGYDLAHLGFPSPSIVREDTLTKPGKIDMKTDWDFYQ 130
RESULT 9
VAV3_MOUSE
ID VAV3_MOUSE STANDARD; PRT; 847 AA.
AC Q9R0C8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE VAV-3 protein.
GN VAV3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
```

SEQUENCE FROM N.A.
Trenkle T., McClelland M., Welsh J.;
Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: EXCHANGE FACTOR FOR GTP-BINDING PROTEINS RHOA, RHOG AND,
TO A LESSER EXTENT, RAC-1. BINDS PHYSICALLY TO THE NUCLEOTIDE-FREE
STATES OF THOSE GTPASES (BY SIMILARITY).
-!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA (SHOWN HERE) AND BETA;
ARE PRODUCED BY ALTERNATIVE SPLICING.
-!- SIMILARITY: CONTAINS 1 CALPONIN-HOMOLOGY (CH) DOMAIN.
-!- SIMILARITY: CONTAINS 1 DBL-HOMOLOGY DOMAIN (DH).
-!- SIMILARITY: CONTAINS 1 PH DOMAIN.
-!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
BINDING DOMAIN.
-!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
-!- SIMILARITY: CONTAINS 2 SH3 DOMAINS.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF067816; AAF09171.1; -
HSSP; PI2931; 1AIE.
MGD; MGI:1888518; Vav3.
InterPro; IPR003247; CH_type.
InterPro; IPR001715; Calponin_hom.
InterPro; IPR002219; DAG_PE_bind.
InterPro; IPR001331; GDS_CDC24.
InterPro; IPR001849; PH.
InterPro; IPR000219; RhoGEF.
InterPro; IPR000980; SH2.
InterPro; IPR001452; SH3.
InterPro; IPR003096; SM22_calponin.
Pfam; PF00307; CH; 1.
Pfam; PF00130; DAG_PE_bind; 1.
Pfam; PF00169; PH; 1.
Pfam; PF00621; RhoGEF; 1.
Pfam; PF00017; SH2; 1.
Pfam; PF00018; SH3; 1.
PRINTS; PR00401; SH2DOMAIN.
PRINTS; PR00452; SH3DOMAIN.
PRINTS; PR00888; SM22CALPONIN.
ProDom; PD001527; CH_type; 1.
SMART; SM00109; C1; 1.
SMART; SM00033; CH; 1.
SMART; SM00233; PH; 1.
SMART; SM00323; RhoGEF; 1.
SMART; SM00252; SH2; 1.
SMART; SM00326; SH3; 2.
PROSITE; PS50021; CH; 1.
PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
PROSITE; PS50010; DH_2; 1.
PROSITE; PS00741; DH_1; 1.
PROSITE; PS50003; PH_DOMAIN; 1.
PROSITE; PS50001; SH2; 1.
PROSITE; PS50002; SH3; 2.
Phorbol-ester binding; Zinc; SH2 domain; SH3 domain; Repeat;
Guanine-nucleotide releasing factor.
DOMAIN 1 119
FT DOMAIN 192 371
FT DOMAIN 400 502
FT DOMAIN 514 562
FT DOMAIN 552 660
FT DOMAIN 672 766
FT DOMAIN 788 847
SQ SEQUENCE 847 AA; 97946 MW; 9AGB63F0D9E60F8F CRC64;

Query Match

15.0%; Score 88.5; DB 1; Length 847;

Best Local Similarity 31.3%; Pred. No. 0.14;
Matches 21; Conservative 14; Mismatches 21; Indels 11; Gaps 2;
QY 23 ISLASAQEDYNAPDCRFINVKKQQIIVVSKLVKENGACGEFWAGSVYGDQDENGVGVF 82
DB 790 LGIATARYPCARDMRELSLLKGMVKIYTKM-----SANGWWRGEVNGR-----VGWF 838
QY 83 PRNLVKE 89
DB 839 PSTYVEE 845
RESULT 10
ID VAV_HUMAN STANDARD; PRT; 845 AA.
AC P15198; Q15860;
DT 01-APR-1990 (Rel. 14, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE VAV proto-oncogene.
GN VAV1 OR VAV.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA DenKinger D.J., Borges C.R., Shaw C.L., Cushman A.M., Kawahara R.S.;
RT "Transcriptional regulation of the vav proto-oncogene."
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 62-845 FROM N.A.
RX MEDLINE=90005432; PubMed=2477241;
RA Katzav S., Martin-Zanca D., Barbacid M.;
RT "vav, a novel human oncogene derived from a locus ubiquitously
expressed in hematopoietic cells."
RL EMBO J. 8:2283-2290(1989).
RN [3]
RP SEQUENCE OF 1-61 FROM N.A.
RX MEDLINE=91172176; PubMed=2005887;
RA Katzav S., Cleveland J.L., Heslop H.E., Pulido D.;
RT "Loss of the amino-terminal helix-loop-helix domain of the vav proto-
oncogene activates its transforming potential."
RL Mol. Cell. Biol. 11:1912-1920(1991).
RN [4]
RP SEQUENCE OF 299-837 FROM N.A.
RA Romero F.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 299-334 FROM N.A.
RX MEDLINE=96038895; PubMed=7478592;
RA Ramos-Norales F., Romero F., Schweighoffer F., Bismuth G., Camonis J.,
RA Tortolero M., Fischer S.;
RT "The proline-rich region of Vav binds to Grb2 and Grb3-3."
RL Oncogene 11:1665-1669(1995).
RN [6]
RP SIMILARITY TO CDC24 FAMILY.
RX MEDLINE=92228488; PubMed=1565462;
RA Adams J.M., Houston H., Allen J., Lints T., Harvey R.;
RT "The hematopoietically expressed vav proto-oncogene shares homology
with the db1 GDP-GTP exchange factor, the bcr gene and a yeast gene
(CDC24) involved in cytoskeletal organization."
RL Oncogene 7:611-618(1992).
CC -!- FUNCTION: PROBABLE EXCHANGE FACTOR FOR A SMALL RAS-LIKE GTP-
BINDING PROTEIN. CAN BE ACTIVATED BY TRUNCATION OF THE N-TERMINUS.
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN HEMATOPOIETIC CELLS BUT
NOT IN OTHER CELL TYPES.
CC -!- SIMILARITY: CONTAINS 1 CALPONIN-HOMOLOGY (CH) DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 DBL-HOMOLOGY DOMAIN (DH).
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
BINDING DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

DE VAV-2 protein.
GN VAV2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=1090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=96313271; PubMed=8710375;
RA Schuebel K.E., Bustelo X.R., Nielsen D.A., Song B.J., Barbacid M.,
RA Goldman D., Lee I.J.;
RT Isolation and characterization of murine vav2, a member of the vav
RT family of proto-oncogenes.;
RL Oncogene 13:363-371(1996).
CC -!- FUNCTION: PROBABLE EXCHANGE FACTOR FOR A SMALL RAS-LIKE GTP-
CC BINDING PROTEIN.
CC -!- SIMILARITY: CONTAINS 1 CALPONIN-HOMOLOGY (CH) DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 DBL-HOMOLOGY DOMAIN (DH).
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC BINDING DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U37017; AAC52761.1; .
DR HSSP; G06631; IGBO.
DR MGD; MGI:102718; Vav2.
DR InterPro; IPR003247; CH_type.
DR InterPro; IPR001715; Calponin_hom.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR001331; GDS_CDC24.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000219; RhoGEF.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00307; CH; 1.
DR Pfam; PF00130; DAG_PE-bind; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00621; RhoGEF; 1.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 2.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD001527; CH_type; 1.
DR SMART; SM00109; C1; 1.
DR SMART; SM00033; CH; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00325; RhoGEF; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS00021; CH; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS00010; DH_2; 1.
DR PROSITE; PS00741; DH_1; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 2.
DR Phosphor-ester binding; zinc; SH2 domain; SH3 domain;
DR Guanine-nucleotide releasing factor; Repeat.
FT DOMAIN 1 119
FT DOMAIN 193 371
FT DOMAIN 400 502
FT DOMAIN 514 562
FT DOMAIN 576 642
SH3 1.

FT DOMAIN 663 757 SH2.
FT DOMAIN 806 867 SH3 2.
SQ SEQUENCE 868 AA; 99915 MW; D18581E7EBB2DBC2 CRC64;

Query Match 14.0%; Score 82.5; DB 1; Length 868;
Best Local Similarity 27.6%; Pred. NO. 0.61;
Matches 21; Conservative 20; Mismatches 22; Indels 13; Gaps 3;

QY 20 VVT---ISLASAQEDYNAPDCRFINVKKQQIYVYVSKLVKENGAGEFWAGSVYGGQDEM 76
Db 802 VFTPRVIGTAVARYNFAARDMRELSLRGDDVVKYISRIIGDQG---WMKG-----ETN 851
QY 77 GVVGYPFRNLVKEQVR 92
Db 852 GRIGWFPSTYYEEGV 867

RESULT 15
VAV2_HUMAN
ID VAV2_HUMAN STANDARD; PRT; 878 AA.
AC P52735;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE VAV-2 protein.
GN VAV2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95283235; PubMed=7762982;
RA Henske E.P., Short M.P., Jozwiak S., Bovey C.M., Ramlakhan S.,
RA Haines J.L., Kwiatkowski D.J.;
RT "Identification of VAV2 on 9q34 and its exclusion as the tuberous
RT sclerosis gene TSC1.";
RL Ann. Hum. Genet. 59:25-37(1995).
CC -!- FUNCTION: PROBABLE EXCHANGE FACTOR FOR A SMALL RAS-LIKE GTP-
CC BINDING PROTEIN.
CC -!- SIMILARITY: CONTAINS 1 CALPONIN-HOMOLOGY (CH) DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 DBL-HOMOLOGY DOMAIN (DH).
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC BINDING DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S76992; AAB34377.1; .
DR HSSP; P08631; IBUL.
DR MIM; 600428; .
DR InterPro; IPR003247; CH_type.
DR InterPro; IPR001715; Calponin_hom.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR001331; GDS_CDC24.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000219; RhoGEF.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00307; CH; 1.
DR Pfam; PF00130; DAG_PE-bind; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00621; RhoGEF; 1.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 2.

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2002, 14:51:06 ; Search time 49.48 Seconds
(without alignments)
1638.219 Million cell updates/sec

Title: US-10-019-455A-23
Perfect score: 330
Sequence: 1 catggaatatttgacgcg.....atattgacttcttcgcgag 330

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

al number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Score	Query Match	Length	DB	ID	Description			
1	63.4	19.2	459	1	US-08-578-649-1	Sequence 1, Appli			
2	63.2	19.2	581	1	US-08-578-649-4	Sequence 4, Appli			
3	57.4	17.4	330	1	US-08-578-649-18	Sequence 18, Appl			
4	54.2	16.4	305	1	US-08-578-649-8	Sequence 8, Appli			
5	40.4	12.2	596	1	US-08-578-649-24	Sequence 24, Appl			
6	40.4	12.2	3565	1	US-08-578-649-3	Sequence 3, Appli			
C 7	34	10.3	7218	1	US-08-232-463-14	Sequence 14, Appl			
8	33.6	10.2	1136	4	US-08-860-820-1	Sequence 1, Appli			
C 9	33.6	10.2	1929	4	US-09-359-161-4	Sequence 4, Appli			
C 10	33.6	10.2	2159	3	US-08-286-870A-7	Sequence 7, Appli			
11	31.2	9.5	4403765	4	US-09-103-840A-2	Sequence 2, Appl			
12	31.2	9.5	4411529	4	US-09-103-840A-1	Sequence 1, Appl			
C 13	30.8	9.3	1736	4	US-09-360-197-13	Sequence 13, Appl			
14	30.4	9.2	289	4	US-09-007-005-17	Sequence 17, Appl			
15	30.4	9.2	289	4	US-09-244-796-17	Sequence 17, Appl			
C 16	29.6	9.0	1000	4	US-09-018-584A-38	Sequence 38, Appl			
17	29.4	8.9	2718	4	US-09-651-656-14	Sequence 14, Appl			
18	29.2	8.8	3101	2	US-08-868-786-1	Sequence 1, Appli			
C 19	28.6	8.7	1518	2	US-08-929-501-4	Sequence 4, Appli			
C 20	28.6	8.7	1518	2	US-08-929-501-5	Sequence 5, Appli			
C 21	28.6	8.7	1518	3	US-09-140-177-4	Sequence 4, Appli			
C 22	28.6	8.7	1518	3	US-09-140-177-5	Sequence 5, Appli			
C 23	28.6	8.7	1518	4	US-09-397-979-4	Sequence 4, Appli			
C 24	28.6	8.7	1518	4	US-09-397-979-5	Sequence 5, Appli			
C 25	28.6	8.7	1753	2	US-08-929-501-1	Sequence 1, Appli			
C 26	28.6	8.7	1753	2	US-08-929-501-3	Sequence 3, Appli			
C 27	28.6	8.7	1753	3	US-09-140-177-1	Sequence 1, Appli			

28	28.6	8.7	1753	3	US-09-140-177-3	Sequence 3, Appli
C 29	28.6	8.7	1753	4	US-09-397-979-1	Sequence 1, Appli
C 30	28.6	8.7	1753	4	US-09-397-979-3	Sequence 3, Appli
C 31	28.6	8.7	1770	2	US-08-929-501-11	Sequence 11, Appl
C 32	28.6	8.7	1770	2	US-08-929-501-13	Sequence 13, Appl
C 33	28.6	8.7	1770	3	US-09-140-177-11	Sequence 11, Appl
C 34	28.6	8.7	1770	3	US-09-140-177-13	Sequence 13, Appl
C 35	28.6	8.7	1770	4	US-09-397-979-11	Sequence 11, Appl
C 36	28.6	8.7	1770	4	US-09-397-979-13	Sequence 13, Appl
C 37	28.6	8.7	2073	2	US-08-929-501-25	Sequence 25, Appl
C 38	28.6	8.7	2073	2	US-08-929-501-26	Sequence 26, Appl
C 39	28.6	8.7	2073	3	US-09-140-177-25	Sequence 25, Appl
C 40	28.6	8.7	2073	3	US-09-140-177-26	Sequence 26, Appl
C 41	28.6	8.7	2073	4	US-09-397-979-25	Sequence 25, Appl
C 42	28.6	8.7	2073	4	US-09-397-979-26	Sequence 26, Appl
C 43	28.6	8.7	2153	4	US-09-367-206-6	Sequence 6, Appli
C 44	28.6	8.7	2576	3	US-09-265-108-1	Sequence 1, Appli
C 45	28.6	8.7	2576	4	US-09-479-264-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-578-649-1
; Sequence 1, Application US/08578649
; Patent No. 5770366
; GENERAL INFORMATION:
; APPLICANT: Ulrich Bogdan
; APPLICANT: Reinhard Buttner
; APPLICANT: Brigitte Kaluza
; TITLE OF INVENTION: MELANOMA-INHIBITING PROTEIN
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578,649
; FILING DATE: 29-July-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 43 24 247.2
; FILING DATE: 20-July-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Andrew L. Tiajolloff
; REGISTRATION NUMBER: 31,575
; REFERENCE/DOCKET NUMBER: BOER 1035-PPE/ALT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; NAME/KEY: CDS
; LOCATION: 40..432
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 40..111
; FEATURE:


```
Query Match      12.2%; Score 40.4; DB 1; Length 596;
Best Local Similarity 63.3%; Pred. No. 0.00087;
Matches 62; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
```

```

RESULT 6
US-08-578-649-3
; Sequence 3, Application US/08578649
; Patent No. 5770366
; GENERAL INFORMATION:
; APPLICANT: Ulrich Bogdan
; APPLICANT: Reinhard Buttner
; APPLICANT: Brigitte Kaluza
; TITLE OF INVENTION: MELANOMA-INHIBITING PROTEIN
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: -805 Third Avenue
; CITY: New York

```

	Query Match	12.2%	Score 40.4;	DB 1;	Length 3565;
	Best Local Similarity	63.3%;	Pred. No.	0.002;	
	Matches	62; Conservative	0; Mismatches	36; Indels	0; Gaps
Qy	65 ctattctcggctagtgctcaagaagtataatgcccgaactgttagattcataacg	124			
pb	1598 ctatcccttcaccgccttcccccccccccgccgacgccgcatccgacgacctcaccca	1648			

	Query Match	12.2%	Score 40.4;	DB 1;	Length 3565;
	Best Local Similarity	63.3%;	Pred. No. 0.002;		
	Matches	62;	Conservative	0;	Mismatches 36; Indels 0; Gaps 0;
Qy	65	ctatttctctggtagtgctcaagaagattataatgcccggaactgtagattcattaacg	124		
Db	1589	CTATCTCATGGCTGTGGCCCTTCAGGACTACATGGCCCGCACTGCCGATTCTTGACCA	1648		
Qy	125	ttaaaaaagggcagacagtctatgtgtactcaaaagctg	162		
Db	1649	TTTACCGGGGCCAAGTGGTGTGATGTCTTCTCCAAGCTG	1686		

RESULT 7
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHETTLINGER, F.


```
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1929
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; OTHER INFORMATION: Lycopersicon esculentum plant homolog of yeast
; OTHER INFORMATION: SNE1 kinase subunit of protein kinase (LeSNF1)
US-09-359-161-4

Query Match
Best Local Similarity 10.2%; Score 33.6; DB 4; Length 1929;
Matches 72; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 84 tcaagagattataatgccccggactgtagattcattacgttaaaagggcagagat 143
Db 1793 TCTGTAAGTAGACTAGGTAACTGAAATTCATATAACGTTAAAAAGAGAGAGCT 1734
144 ctatgttactcaagctgttaaaagaaatggagctggagaaattttggctggcagtgt 203
Db 1733 AGATCTCCTAGCGAATCAAAAGAAAAAGAGTTGCTAGGCACCTCTGACCAAGAGGGT 1674

QY 204 ttatggtgatggccag 219
Db 1673 GGAATTTGTTAAACAAG 1658

RESULT 10
US-08-286-870A-7/c
; Sequence 7, Application US/08286870A
; Patent No. 6063605
; GENERAL INFORMATION:
; APPLICANT: ELY, S
; APPLICANT: TAILOR, RH
; APPLICANT: TIPPETT, JM
; APPLICANT: BLENK, RG
; TITLE OF INVENTION: BACTERIAL GENES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABBY & CUSHMAN
; ADDRESSEE: Intellectual Property Group of
; ADDRESSEE: PILLSBURY, MADISON & SUTRO LLP
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,870A
; FILING DATE: 05-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/520228
; FILING DATE: 09-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8910624.9
; FILING DATE: 09-MAY-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: PAUL N. KOKULIS
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 70608/220720
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 2159 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2159
US-08-286-870A-7

Query Match
Best Local Similarity 10.2%; Score 33.6; DB 3; Length 2159;
Matches 75; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 174 tggagctggagaattttggctggcagtggtttatgtgtatggccagagagatgggagt 233
Db 2002 TGGTGTGTGAACAGGGCGGTGACCTTCCTCGGGCCTTCTCGAAGTCGTACTCGGCT 1943

QY 234 cgtgggttattttcccccaggaacttggtcaagggaacagcgtgtgtaccaggaagctaccaa 293
Db 1942 CGTAGTCACTCCACGGGCGACCAACTCGATGCGGTGATGATACACTCGTTCCCGCTGC 1883

QY 294 ggaagttcccccaccaggtatttga 317
Db 1882 TGAAGTTCAGGCGCGCATGGTGA 1859

RESULT 11
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match
Best Local Similarity 9.5%; Score 31.2; DB 4; Length 4403765;
Matches 90; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 99 tgcgccggactgtagattcatttaacgttaaaaaagggcagcagatctatgtgtactcaaa 158
Db 426225 tgtccagcgggcccgcgcgtcccccgcattgcagggtgatcgaccgacggggccaaga 426284

QY 159 gctgttaaaagaaatggagcgtggagaattttggcgtgcagctgttatgtgtatggcca 218
Db 426285 cgggtccactgaggatgcctatgtcgacgcttcgcgtggcgtatctctggggaacgtaa 426344

QY 219 ggacgagatgggagtcgtgtgtttatttccccaggaactgggtcaagaaacagcgtgtgta 278
Db 426345 tggcggggatggagacatgtgtatgtcgcggtgatggggaagtggtgacccgggatgtcga 426404

QY 279 ccagggaag 286
```


Db 426405 cattgagg 426412

RESULT 12

US-09-103-840A-1

; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Query Match 9.5%; Score 31.2; DB 4; Length 4411529;
Best Local Similarity 47.9%; Pred. No. 34;
Matches 90; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 99 tgcctcggagctagattcattaaacgttaaaagggcagcagatctgtgtactcaaa 158
Db 426067 tgcctcagcggggccagcgtcccgccatgcaggtgtatcgaccggggccaaaga 426126
QY 159 gctgttaaaagaaatggagctggagaattttggcgtggcagtggtttatggatggcca 218
Db 426127 cgggtccactgaggatgctatgtcagcgtccctcgtcgtcgtatcttggggaacgttaa 426186
QY 219 ggacgagatgggctgctgggttatttccccaggaaacttggtcaaggaaacgctgtgtga 278
Db 426187 tggcgggagtgagacattggtgctgcgcggtgtggtggtggtgacccggatgtcga 426246
QY 279 ccaggaag 286
Db 426247 cattgagg 426254

RESULT 13

US-360-197-13/C

; Sequence 13, Application US/09360197
; Patent No. 6287859
; GENERAL INFORMATION:
; APPLICANT: Bassilana, Frederic
; APPLICANT: Lazdunski, Michel
; APPLICANT: Waldmann, Rainer
; APPLICANT: Deweille, Jan R.
; TITLE OF INVENTION: Human and Rat Families of Neuronal Acid-Sensitive
; TITLE OF INVENTION: Cationic Channels, Their Cloning and Applications
; FILE REFERENCE: 989.6706P
; CURRENT APPLICATION NUMBER: US/09/360.197
; CURRENT FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: 09/129,758
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/095,408
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1736
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-360-197-13

Query Match 9.3%; Score 30.8; DB 4; Length 1736;
Best Local Similarity 63.5%; Pred. No. 1.8;
Matches 47; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 247 ccaggaacttggtcaaggaacagcgtgtgtaccaggaagactaccaaggaagtccacc 306
Db 533 CCCACAAGGTTGGCCACCGAAGCAGACAGTCCAGCAGCATGTATCCAGGAGTGCCCGC 474
QY 307 acggatattgactt 320
Db 473 ACGGGCATAGATT 460

RESULT 14

US-09-007-005-17
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007.005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

Query Match 9.2%; Score 30.4; DB 4; Length 289;
Best Local Similarity 6.6%; Pred. No. 1.1;
Matches 16; Conservative 96; Mismatches 130; Indels 0; Gaps 0;

QY 56 gtgtctactatttcttctgtgtgtctcaagaagattataatgcccgagcttagat 115
Db 27 rurururcrarururcrarururgrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnr 86
QY 116 tcattaacgttaaaaaaggcagcagatctgtgtactcaagctggtaaaagaaaatg 175
Db 87 rnrrnr 146
QY 176 gagctggagaatttggcgtggcagtggtttatggtgatggccagcagagatggagtcg 235
Db 147 rnrrnr 206
QY 236 tgggttatttccccaggaacttggtcaagaaacagcgtgtgtaccaggaagctaccaag 295
Db 207 rnrrnr 266
QY 296 aa 297
Db 267 aa 268

RESULT 15

```

US-09-244-796-17
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 09/007,005
; EARLIER FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 289
TYPE: RNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Translation template
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(289)
OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

```

```

Query Match          9.2%; Score 30.4; DB 4; Length 289;
Best Local Similarity 6.6%; Pred. No. 1.1;
Matches 16; Conservative 96; Mismatches 130; Indels 0; Gaps 0;

QY 56 gtgtctactattctctgtgtagtgcacgaagattataatgcccgagctgtagat 115
   : : : : | : : : : | : : : : | : : : : | : : : : | : : : :
Db 27 rururarcrarururarcrarurgrnrnrnrnrnrnrnrnrnrnrnrnrnr 86

QY 116 tcattaacgttaaaaaaggcgacgagatctatgttactcaagctggtaaaaaatg 175
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 87 rnrrsrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnr 146

QY 176 gagctggagaaatttggcgtggcagtttattgtggccagcagagatgggagtcg 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 147 rnrrsrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnr 206

QY 236 tgggttattcccgaggaaactgtgtcagggaacagcgtgtgtaccaggaagctaccaagg 295
   : : : : : : : : : : | : : : : | : : : : | : : : : | : : : :
Db 207 rnrrsrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnrnr 266

QY 296 aa 297
   ||
Db 267 aa 268

```

Search completed: September 23, 2002, 15:39:48
Job time: 2922 sec

SHINICHI MOGI,
PI YUKO NOGUCHI, KOJI YOSHIMURA, HIDEYUKI TANAKA
PC C12N15/09, A61K38/00, A61K45/00, A61K48/00, A61P9/00, A61P19/02, PC
A61P19/08,
PC C07K14/47, C07K16/18, C12N1/21, C12N5/10, G01N33/15, G01N33/50, PC
G01N33/53//
PC C12P21/08, C12N15/00, A61K37/02, C12N5/00
CC

CC Key Location/Qualifiers
FH source 1..330
FT /organism='Homo sapiens (human)'.
FT Location/Qualifiers

FEATURES

source

/organism='Homo sapiens'
/db_xref='taxon:9606'
91 a 60 c 91 g 88 t

BASE COUNT
ORIGIN

Query Match 100.0%; Score 330; DB 6; Length 330;

Best Local Similarity 100.0%; Pred. No. 3e-88;

Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 catggaattattatgaccgtctagcttccaagaagctctgtgcaagatgatgagtgtgc 60

Db 1 CATGGAATATTATGACCCTCTAGCTTCCAAGAAGCTCTGTGCAGATGATGAGTGTGC 60

QY 61 tatactatttctgtgctagctcgaagattataatgccccggactgtgattcatt 120

Db 61 TATACTATTCTCTGCTAGTCTCAAGAAGATTATAATGCCCCGACGTGATATTCATT 120

QY 121 aacgttaaaaggcagcagatctctgtactcaaaagctggttaaaagaatggagct 180

Db 121 AACGTTAAAAAGGCGAGCATCTATGTGTACTCAAGCTGGTAAAGAAAAATGGAGCT 180

QY 181 gagaattttggctgcagcttttatgtgatggccagcagagatggagtcgtggt 240

Db 181 GGAAATTTTGGCTGCGAGTGTATTATGTTATGCGCAGAGATGGAGTGTGCGGT 240

QY 241 tatttccccaggaacttggcgaagacagcgtgtgtaccaggaagctaccaaggaagtt 300

Db 241 TATTTCGCCAGGAACCTTGGTCAAGGAACAGCGTGTGTACCAGGAAGCTACCAAGGAAGTT 300

QY 301 cccaccacggatattgaactttcttcgcgag 330

Db 301 CCCACCACGGATATTGACTTCTCTCCGAG 330

RESULT 2

LOCUS

BD010802 Novel polypeptide and DNA thereof.

ACCESSION BD010802

VERSION BD010802.1 GI:18639175

KEYWORDS JP 2001069994-A/3.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 384)

Ito, Y., Nishi, K., Ogi, K., Okubo, S., Mogi, S., Noguchi, Y.,

Yoshimura, K. and Tanaka, H.

Novel polypeptide and DNA thereof

Patent: JP 2001069994-A 3 21-MAR-2001;

TAKEDA CHEMICAL INDUSTRIES LTD

OS Homo sapiens (human)

PN JP 2001069994-A/3

PD 21-MAR-2001

PF 29-JUN-2000 JP 2000195911

PI YASUAKI ITO, KAZUNORI NISHI, KAZUHIRO OGI, SHOICHI OKUBO, PI

SHINICHI MOGI,

PI YUKO NOGUCHI, KOJI YOSHIMURA, HIDEYUKI TANAKA

PC C12N15/09, A61K38/00, A61K45/00, A61K48/00, A61P9/00, A61P19/02, PC
A61P19/08,
PC C07K14/47, C07K16/18, C12N1/21, C12N5/10, G01N33/15, G01N33/50, PC
G01N33/53//
PC C12P21/08, C12N15/00, A61K37/02, C12N5/00
CC

CC Key Location/Qualifiers
FH source 1..384
FT /organism='Homo sapiens (human)'.
FT Location/Qualifiers

FEATURES

source

/organism='Homo sapiens'
/db_xref='taxon:9606'
99 a 70 c 106 g 109 t

BASE COUNT
ORIGIN

Query Match 100.0%; Score 330; DB 6; Length 384;

Best Local Similarity 100.0%; Pred. No. 3.1e-88;

Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 catggaattattatgaccgtctagcttccaagaagctctgtgcaagatgatgagtgtgc 60

Db 55 CATGGAATATTATGACCCTCTAGCTTCCAAGAAGCTCTGTGCAGATGATGAGTGTGC 114

QY 61 tatactatttctgtgctagctcgaagattataatgccccggactgtgattcatt 120

Db 115 TATACTATTCTCTGCTAGTCTCAAGAAGATTATAATGCCCCGACGTGATATTCATT 174

QY 121 aacgttaaaaggcagcagatctctgtactcaaaagctggttaaaagaatggagct 180

Db 175 AACGTTAAAAAGGCGAGCATCTATGTGTACTCAAGCTGGTAAAGAAAAATGGAGCT 234

QY 181 gagaattttggctgcagcttttatgtgatggccagcagagatggagtcgtggt 240

Db 235 GGAAATTTTGGCTGCGAGTGTATTATGTTATGCGCAGAGATGGAGTGTGCGGT 294

QY 241 tatttccccaggaacttggcgaagacagcgtgtgtaccaggaagctaccaaggaagtt 300

Db 295 TATTTCGCCAGGAACCTTGGTCAAGGAACAGCGTGTGTACCAGGAAGCTACCAAGGAAGTT 354

QY 301 cccaccacggatattgaactttcttcgcgag 330

Db 355 CCCACCACGGATATTGACTTCTCTCCGAG 384

RESULT 3

LOCUS

AF233261 Homo sapiens otoraplin (OTOR) mRNA, complete cds.

ACCESSION AF233261

VERSION AF233261.1 GI:8927427

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 846)

Robertson, N.G., Heller, S., Lin, J.S., Resendes, B.L., Weremowicz, S.,

Denis, C.S., Bell, A.M., Hudspeth, A.J., and Morton, C.C.

A Novel Conserved Cochlear Gene, OTOR: Identification, Expression

Analysis, and Chromosomal Mapping

Genomics 66 (3), 242-248 (2000)

10873378

2 (bases 1 to 846)

Robertson, N.G., Heller, S., Lin, J.S., Resendes, B.L., Weremowicz, S.,

Denis, C.S., Bell, A.M., Hudspeth, A.J., and Morton, C.C.

Direct Submission

Submitted (09-FEB-2000) Pathology, Brigham and Women's Hospital, 75

Francis Street, Boston, MA 02115, USA

Location/Qualifiers

1..846

/organism='Homo sapiens'

```

/db_xref="taxon:9606"
/chromosome="20"
/map="20p11.23-p12.1"
/tissue_type="cochlea"
/dev_stage="fetus"
1. 846
/gene="OTOR"
45. 431
/gene="OTOR"
/note="similar to cdrap/mia"
/codon_start=1
/product="otoraplin"
/protein_id="AAF82078.1"
/db_xref="GI:8927428"
/translation="MARILLLEPLGLVAVCAVHGIFMDRLASKKLCADDECVYTISLA
SAQEDNAPDCRFINVKQQIYVYSLVKENGAGEFWAGSVYGDQEMGVGVFPR
NLVKQRYVQEAETKEVPTDIDFCE"
BASE COUNT 235 a 147 c 205 g 259 t
ORIGIN

Query Match 100.0%; Score 330; DB 9; Length 846;
Best Local Similarity 100.0%; Pred. No. 3.4e-88;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 catggaattatttgaccgtctagcttccaaagaagctctgtgcagatgatgagtgctc 60
|||||
Db 99 CATGGAATATTTATGACCGCTCTAGCTTCCAAGAAGCTCTGTGCAGATGATGAGTGTGTC 158
|||||
QY 61 tatactatttctgtgctgctcaagaagattataatgcccgactgtagattcatt 120
|||||
Db 159 TATACATATTTCTGCTAGTGTCTCAAGAAGATTATAATGCCCGACTGTAGATTCATT 218
|||||
QY 121 aacgttcaaaaaggcgagcagatctatgtgtactcaaaagctgtgtaaaagaaatggagct 180
|||||
Db 219 AACGTTAAAAAGGCGCAGCAGATCTATGTGTACTCAAAAGCTGTAAGAAAAATGGAGCT 278
|||||
QY 181 ggagaatttggctggcagtgcttattgtatgtgacagacagatggagtcgtgggt 240
|||||
Db 279 GGAGAAATTTGGCTGGCAGTGTATGTTATGTTGATGGCCAGCAGATGGAGTGGTGGGT 338
|||||
QY 241 tatttccccaggaacttggctcaaggacagcgtgtgtaccaggaagctaccagaaggaatt 300
|||||
Db 339 TATTTCCCGAGGAACCTGGTCAAGGAACAGCGTGTGTACCAGGAAGCTACCAAGGAAGTT 398
|||||
QY 301 cccacacggatattgacttcttctgcgag 330
|||||
Db 399 CCCACCACGGATATTGACTTCTTCTGCGAG 428

RESULT 4
AF243505 AF243505 865 bp mRNA linear PRI 26-DEC-2000
LOCUS Homo sapiens fibrocyte-derived protein (FDP) mRNA, complete cds.
DEFINITION AF243505
ACCESSION AF243505
VERSION AF243505.1 GI:11991843
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 865)
AUTHORS Cohen-Salmon, W., Frensch, D., Liu, W., Verpy, E., Voegelings, S. and
Petit, C.
TITLE Fdp, a New Fibrocyte-derived Protein Related to MIA/CD-RAP, Has an
in Vitro Effect on the Early Differentiation of the Inner Ear
Mesenchyme
J. Biol. Chem. 275 (51), 40036-40041 (2000)
JOURNAL 10998416
PUBMED 2 (bases 1 to 865)
REFERENCE Cohen-Salmon, M., Frensch, D., Verpy, E., Voegelings, S. and Petit, C.
AUTHORS Direct Submission
TITLE Submitted (09-MAR-2000) Biotechnologies, Institut Pasteur, 25 rue
JOURNAL

```

```

du Dr. Roux, Paris 75015, France
Location/Qualifiers
1. 865
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p11"
1. 865
/gene="FDP"
8. 394
/gene="FDP"
/codon_start=1
/product="fibrocyte-derived protein"
/protein_id="AAG42356.1"
/db_xref="GI:11991844"
/translation="MARILLLEPLGLVAVCAVHGIFMDRLASKKLCADDECVYTISLA
SAQEDNAPDCRFINVKQQIYVYSLVKENGAGEFWAGSVYGDQEMGVGVFPR
NLVKQRYVQEAETKEVPTDIDFCE"
BASE COUNT 258 a 145 c 207 g 255 t
ORIGIN

Query Match 100.0%; Score 330; DB 9; Length 865;
Best Local Similarity 100.0%; Pred. No. 3.4e-88;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 catggaattatttgaccgtctagcttccaaagaagctctgtgcagatgatgagtgctc 60
|||||
Db 62 CATGGAATATTTATGACCGCTCTAGCTTCCAAGAAGCTCTGTGCAGATGATGAGTGTGTC 121
|||||
QY 61 tatactatttctgtgctgctcaagaagattataatgcccgactgtagattcatt 120
|||||
Db 122 TATACATATTTCTGCTAGTGTCTCAAGAAGATTATAATGCCCGACTGTAGATTCATT 181
|||||
QY 121 aacgttcaaaaaggcgagcagatctatgtgtactcaaaagctgtgtaaaagaaatggagct 180
|||||
Db 182 AACGTTAAAAAGGCGCAGCAGATCTATGTGTACTCAAAAGCTGTAAGAAAAATGGAGCT 241
|||||
QY 181 ggagaatttggctggcagtgcttattgtatgtgacagacagatggagtcgtgggt 240
|||||
Db 242 GGAGAAATTTGGCTGGCAGTGTATGTTATGTTGATGGCCAGCAGATGGAGTGGTGGGT 301
|||||
QY 241 tatttccccaggaacttggctcaaggacagcgtgtgtaccaggaagctaccagaaggaatt 300
|||||
Db 302 TATTTCCCGAGGAACCTGGTCAAGGAACAGCGTGTGTACCAGGAAGCTACCAAGGAAGTT 361
|||||
QY 301 cccacacggatattgacttcttctgcgag 330
|||||
Db 362 CCCACCACGGATATTGACTTCTTCTGCGAG 391

RESULT 5
BD010820 BD010820 923 bp DNA linear PAT 31-JAN-2002
LOCUS Novel polypeptide and DNA thereof.
DEFINITION BD010820
ACCESSION BD010820
VERSION BD010820.1 GI:18639193
KEYWORDS JP 2001069994-A/21.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 923)
AUTHORS Ito, Y., Nishi, K., Ogi, K., Okubo, S., Mogi, S., Noguchi, Y.,
Yoshimura, K. and Tanaka, H.
TITLE Novel polypeptide and DNA thereof
JOURNAL Patent: JP 2001069994-A 21 21-MAR-2001;
TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)
PN JP 2001069994-A/21
PD 21-MAR-2001
PF 29-JUN-2000 JP 2000195911
PR

```

PI YASUAKI ITO,KAZUNORI NISHI,KAZUHIRO OGI,SHOICHI OKUBO, PI
SHINICHI MOGI,
PI YUKO NOGUCHI,KOJI YOSHIMURA,HIDEYUKI TANAKA
PC C12N15/09,A61K38/00,A61K45/00,A61K48/00,A61P9/00,A61P19/02, PC
A61P19/08,
PC C07K14/47,C07K16/18,C12N1/21,C12N5/10,G01N33/15,G01N33/50, PC
G01N33/53//
PC C12P21/08,C12N15/00,A61K37/02,C12N5/00
CC
FH Key Location/Qualifiers
FT source i. .923
FT /organism='Homo sapiens (human)'.
Location/Qualifiers
1. .923
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 303 a 147 c 213 g 260 t.
ORIGIN

Query Match 100.0%; Score 330; DB 6; Length 923;
Best Local Similarity 100.0%; Pred. No. 3.5e-88;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 catgaatatattatggacgcttagcttccaagaagctctgtgcagatgatgagtgtgc 60
|||||
Db 88 CATGGAATATTTATGACCGCTAGCTTCCAAGAAGCTCTGTGCAGATGATGAGTGTGC 147
|||||

QY 61 tatactattctctgctagtctcaagaagattataatgccccgactgttagattcatt 120
|||||
Db 148 TATACTATTTCTCTGGCTAGTCTCAAGAAGATTATATGCCCCGACTGTAGATTCAAT 207
|||||

QY 121 aecgttaaaaaagggcagcagatctatgtgtactcaagctgtgttaaaagaaatggagct 180
|||||

Db 208 ACGTTAAAAGGCGAGCAGATCTATGTGTACTCAAGCTGGTAAAGAAAATGGAGCT 267
|||||

QY 181 ggagaattttggcgtggcagtgcttattgtgtatggccaggacgagatgggagtcgtgggt 240
|||||

Db 268 GGAGAAATTTGGCGTGGCAGTGTATGTGTATGTCGTCAGGAGGAGATGGGAGTCGTGGGT 327
|||||

QY 241 tatttccccagaacttgctcaagaacagctgtgtaccagaagctaccagaagatt 300
|||||

Db 328 TATTTCCCGAGGAACCTTGGTCAAGGAACAGCGTGTGTACCGAGGAAGCTACCAAGGAAGTT 387
|||||

QY 301 ccaccacgcatattgacttcttctctgcgag 330
|||||

Db 388 CCCACCACGGATATTGACTTCTCTCGGAG 417
|||||

RESULT 6
HSA242552
LOCUS
DEFINITION Homo sapiens mRNA for melanoma inhibitory activity like protein (MIAL gene).
ACCESSION AJ242552
VERSION AJ242552.1 GI:12619172
KEYWORDS melanoma inhibitory activity like protein; Mial gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1422)
Rendtorff,N.D.,, Frodin,M., Attie-Bitach,T., Vekemans,M. and Tommerup,N.
TITLE Identification and Characterization of an Inner Ear-Expressed Human Melanoma Inhibitory Activity (MIA)-like Gene (MIAL) with a Frequent Polymorphism That Abolishes Translation
JOURNAL Genomics 71 (1), 40-52 (2001)
MEDLINE 21100875
REFERENCE 2 (bases 1 to 1422)
Rendtorff,N.D.
AUTHORS Direct Submission
TITLE Submitted (21-MAY-1999) Rendtorff N.D., Department of Medical

Genetics, Institute of Medical Biochemistry and Genetics, Blegdamsvej 3, 2200 Copenhagen N, DENMARK
Revised by author 03-AUG-1999
Related sequences: AJ242939 (Mus musculus mRNA)
Related sequences: AJ252324 to AJ252327 (genomic sequence).
Location/Qualifiers
1. .1422
/organism='Homo sapiens'
/db_xref='taxon:9606'
/chromosome='20'
/map='20p11.22-12.2'
/tissue_type='brain and cochlea'
/dev_stage='fetal'
<1. .115
/gene='MIAL'
/number=1
1. .387
/gene='MIAL'
/codon_start=1
/product='melanoma inhibitory activity like protein'
/protein_id='CAC27443.1'
/db_xref='GI:12619173'
/translation='MARILLLLFLPLVAVCAVHGIFMDRLASKKLCAADDECYYTISLA
SAQEDYNAPDCRFINVRKKQOIYVYSKLKENGAGEFWAGSVYGGQDEMVGVGIFPR
NLVKEQRYIQEATKEVPTTIDIFFCE'
1. .75
/gene='MIAL'
1. .1422
/gene='MIAL'
<1. .1422
/gene='MIAL'
/product='melanoma inhibitory activity like protein'
76. .384
/product='melanoma inhibitory activity like protein'
/gene='MIAL'
/product='melanoma inhibitory activity like protein'
116. .257
/number=2
115~116
/gene='MIAL'
/number=1
258. .363
/gene='MIAL'
/number=3
257~258
/gene='MIAL'
/number=2
364. .387
/gene='MIAL'
/number=4
363~364
/gene='MIAL'
/number=3
388. .1422
/gene='MIAL'
BASE COUNT 417 a 220 c 306 g 479 t
ORIGIN

Query Match 100.0%; Score 330; DB 9; Length 1422;
Best Local Similarity 100.0%; Pred. No. 3.7e-88;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 catggaattattatggacgcttagcttccaagaagctctgtgcagatgatgagtgtgc 60
|||||

Db 55 CATGGAATATTTATGACCGCTAGCTTCCAAGAAGCTCTGTGCAGATGATGAGTGTGC 114
|||||

QY 61 tatactattctctgctagtctcaagaagattataatgccccgactgttagattcatt 120
|||||

Db 115 TATACTATTTCTCTGGCTAGTCTCAAGAAGATTATATGCCCCGACTGTAGATTCAAT 174
|||||

QY 121 aacgttaaaaaagggcagcagatctatgtgtactcaagctgtgttaaaagaaatggagct 180
|||||

```
Db 175 AACGTTAAAAAGGCGACGAGATCTATGTGTAACAGCTGGTAAAGAAATGGAGCT 234
Qy 181 ggagaattttggcgtggcagtgctttatggtgagccaggagagatggagtcgtgggt 240
Db 235 GGAGAAATTTGGCGTGGCAGTGTATGATGATGCCAGGACGAGATGGAGTCTGGGT 294
Qy 241 tatttccccaggaacttggtcaaggaaacagcgtgtgtaccaggaagctaccagaagatt 300
Db 295 TATTTTCCCGAGAACTTGGTCAAGGAACAGCGTGTGTACCAGGAAGCTACCAAGGAAGTT 354
Qy 301 ccacacagcgatttgactctctctgcgag 330
Db 355 CCACACAGGATATTGACTTCTCTCGCAG 384

RESULT 7
BD010817
LOCUS BD010817 330 bp DNA linear PAT 31-JAN-2002
DEFINITION Novel polypeptide and DNA thereof.
ACCESSION BD010817
VERSION BD010817.1 GI:18639190
KEYWORDS JP 2001069994-A/18.
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Ito, Y., Nishi, K., Ogi, K., Okubo, S., Mogi, S., Noguchi, Y.,
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 330)
Yoshimura, K. and Tanaka, H.
TITLE Novel polypeptide and DNA thereof
JOURNAL Patent: JP 2001069994-A 18 21-MAR-2001;
TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Mus sp. (mouse)
PN JP 2001069994-A/18
PD 21-MAR-2001
PF 29-JUN-2000 JP 2000195911
PR
PI YASUAKI ITO, KAZUNORI NISHI, KAZUHIRO OGI, SHOICHI OKUBO, PI
SHINICHI MOGI,
PI YUKO NOGUCHI, KOJI YOSHIMURA, HIDEYUKI TANAKA
PC C12N15/09, A61K38/00, A61K45/00, A61K48/00, A61P9/00, A61P19/02, PC
A61P19/08,
PC C07K14/47, C07K16/18, C12N1/21, C12N5/10, G01N33/15, G01N33/50, PC
G01N33/53//
PC C12P21/08, C12N15/00, A61K37/02, C12N5/00
CC
FH Key Location/Qualifiers
FT source 1. .330
FT /organism="Mus sp. (mouse)".
FEATURES
source
location/Qualifiers
1. .330
/organism="Mus sp."
/db_xref="taxon:10095"
BASE COUNT 91 a 60 c 92 g 87 t
ORIGIN

Query Match 76.4%; Score 252.2; DB 6; Length 330;
Best Local Similarity 85.4%; Pred. No. 6.6e-65;
Matches 281; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 1 catggaattttatggaccgtctagcttccaagaagctctgtgcagatgatgagtgctc 60
Db 1 CATGTGTATTATGGATAAATCTTCTTCTTAAGAAGTTGTGTCCGATGAGGAGTGTGTC 60
Qy 61 tatattcttctgctagtgctcaagaagattataatgccccgactgtagattcatt 120
Db 61 TATATTCTTCTGCGAAGACGAGGAAGATTACAATGCCCGACGTAGGTTCATC 120
Qy 121 aacgttaaaagggcagcagatctatgtgtactcaagctggttaaaagaaatggagct 180
Db 121 GATGTCAAGAAAGGCGACGAGATCTATGTTTACTCAAGCTGGTAAACAAACGGAGCT 180
```

```
Qy 181 ggagaattttggcgtggcagtgctttatggtgagccaggagagatggagtcgtgggt 240
Db 181 GGAGAGATTTTGGCGTGGCAGTGTATGATGATGCCAGGATGAGATGGAAATTTAGGT 240
Qy 241 tatttccccaggaacttggtcaaggaaacagcgtgtgtaccaggaagctaccagaagatt 300
Db 241 TATTTTCCCGAGAACTTGGTGAAGGAGGAGCGTGTATACCAGGAGGCCACCAAGGAGTC 300
Qy 301 ccacacagcgatttgactctctctgcgga 329
Db 301 CCAACACGATATTGACTTCTCTCTGTA 329

RESULT 8
BD010805
LOCUS BD010805 384 bp DNA linear PAT 31-JAN-2002
DEFINITION Novel polypeptide and DNA thereof.
ACCESSION BD010805
VERSION BD010805.1 GI:18639178
KEYWORDS JP 2001069994-A/6.
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Ito, Y., Nishi, K., Ogi, K., Okubo, S., Mogi, S., Noguchi, Y.,
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 384)
Yoshimura, K. and Tanaka, H.
TITLE Novel polypeptide and DNA thereof
JOURNAL Patent: JP 2001069994-A 6 21-MAR-2001;
TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Mus sp. (mouse)
PN JP 2001069994-A/6
PD 21-MAR-2001
PF 29-JUN-2000 JP 2000195911
PR
PI YASUAKI ITO, KAZUNORI NISHI, KAZUHIRO OGI, SHOICHI OKUBO, PI
SHINICHI MOGI,
PI YUKO NOGUCHI, KOJI YOSHIMURA, HIDEYUKI TANAKA
PC C12N15/09, A61K38/00, A61K45/00, A61K48/00, A61P9/00, A61P19/02, PC
A61P19/08,
PC C07K14/47, C07K16/18, C12N1/21, C12N5/10, G01N33/15, G01N33/50, PC
G01N33/53//
PC C12P21/08, C12N15/00, A61K37/02, C12N5/00
CC
FH Key Location/Qualifiers
FT source 1. .384
FT /organism="Mus sp. (mouse)".
FEATURES
source
location/Qualifiers
1. .384
/organism="Mus sp."
/db_xref="taxon:10095"
BASE COUNT 98 a 68 c 111 g 107 t
ORIGIN

Query Match 76.4%; Score 252.2; DB 6; Length 384;
Best Local Similarity 85.4%; Pred. No. 6.7e-65;
Matches 281; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 1 catggaattttatggaccgtctagcttccaagaagctctgtgcagatgatgagtgctc 60
Db 55 CATGTGTATTATGGATAAATCTTCTTCTTAAGAAGTTGTGTCCGATGAGGAGTGTGTC 114
Qy 61 tatattcttctgctagtgctcaagaagattataatgccccgactgtagattcatt 120
Db 115 TATATTCTTCTGCGAAGACGAGGAAGATTACAATGCCCGACGTAGGTTCATC 174
Qy 121 aacgttaaaagggcagcagatctatgtgtactcaagctggttaaaagaaatggagct 180
Db 175 GATGTCAAGAAAGGCGACGAGATCTATGTTTACTCAAGCTGGTAAACAAACGGAGCT 234
Qy 181 ggagaattttggcgtggcagtgctttatggtgagccaggagagatggagtcgtgggt 240
```



```
Db 235 GGAGAGTTTTGGCGTGGCAGTCTTTATGCTGACCCAGGATGAGATGGGAATTGTAGCT 294
Qy 241 tatttccccagaacttgctcaagaacagctgtgtaccaggaagctaccagaagatt 300
Db 295 TATTTCCCGACGACTTGTGTGAAGACGACGCTGTATACCGAGGCGCCACCAAGGAGATC 354
Qy 301 cccacacagcgatattgactctcttcgcga 329
Db 355 CCAACCACGGATATTGACTTCTTCTGTGA 383

RESULT 9
AF243504 AF243504 929 bp mRNA linear ROD 26-DEC-2000
LOCUS Mus musculus fibrocyte-derived protein (Fdp) mRNA, complete cds.
DEFINITION AF243504
ACCESSION AF243504
VERSION AF243504.1 GI:11991841
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 929)
AUTHORS Cohen-Salmon,M., Frenz,D., Liu,W., Verpy,E., Voegelting,S. and
Petit,C.
TITLE Fdp, a New Fibrocyte-derived Protein Related to MIA/CD-RAP, Has an
in Vitro Effect on the Early Differentiation of the Inner Ear
Mesenchyme
JOURNAL J. Biol. Chem. 275 (51), 40036-40041 (2000)
PUBMED 10998416
REFERENCE 2 (bases 1 to 929)
AUTHORS Cohen-Salmon,M., Frenz,D., Verpy,E., Voegelting,S. and Petit,C.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) Biotechnologies, Institut Pasteur, 25 rue
du Dr. Roux, Paris 75015, France
FEATURES
Source Location/Qualifiers
1..929
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/chromosome="2"
gene 1..929
/gene="Fdp"
CDS 9..395
/gene="Fdp"
/codon_start=1
/product="fibrocyte-derived protein"
/protein_id="AAG42355.1"
/db_xref="GI:11991842"
/translation="MARILILLGLVLCAGHGVFMDKLSKKLKADAECEVYTISLA
RAQEDYNAPDCRFIDVKKGOIYVYSKLVTEENGAGEFWAGSVYGDHDEMGIYGFPS
NLVKEQRYQEAETKEIPTDIDFCE"
BASE COUNT 260 a 156 c 220 g 293 t
ORIGIN

Query Match 76.4%; Score 252.2; DB 10; Length 929;
Best Local Similarity 85.4%; Pred. No. 7.6e-65;
Matches 281; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 1 catggaataattatgacgcgtctagcttccaagaagctctgtgcagatgatgagtgctc 60
Db 63 CATGGTGATTTTATGGATAAACTTCTTCTTAAGAAGTTGTGTGCGGATGAGGAGTGTGTC 122
Qy 61 tatactattctctgctagtgctcaagaattataatgcccgagctgtagattcatt 120
Db 123 TATACTATTCTCTGGCAAGACACAGAGATTACATTACATGCCACAGACTGTAGGTTTCATC 182
Qy 121 aacgttaaaaggcagcagatctatgtgtactcaaaagctgtgtaaaagaaaatggagct 180
Db 183 GATGTCACAGAAAGGCGACGAGATCTATGTTTACTCAAGCTGGTTAACAGAAAACGAGCT 242
Qy 181 ggagaattttgggctggcagtgcttctatggtgatggccaggacagatgggagtcgtgggt 240
```

```
Db 243 GGAGAGTTTTGGCGTGGCAGTCTTTATGCTGACCCAGGATGAGATGGGAATTGTAGCT 302
Qy 241 tatttccccagaacttgctcaagaacagctgtgtaccaggaagctaccagaagatt 300
Db 303 TATTTCCCGACGACTTGTGTGAAGACGACGCTGTATACCGAGGCGCCACCAAGGAGATC 362
Qy 301 cccacacagcgatattgactctcttcgcga 329
Db 363 CCAACCACGGATATTGACTTCTTCTGTGA 391

RESULT 10
BD010821 BD010821 947 bp DNA linear PAT 31-JAN-2002
LOCUS Novel polypeptide and DNA thereof.
DEFINITION BD010821
ACCESSION BD010821
VERSION BD010821.1 GI:18639194
KEYWORDS JP 2001069994-N/22.
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 947)
AUTHORS Ito,Y., Nishi,K., Ogi,K., Okubo,S., Mogi,S., Noguchi,Y.,
Yoshimura,K. and Tanaka,H.
TITLE Novel polypeptide and DNA thereof
JOURNAL Patent: JP 2001069994-A 22 21-MAR-2001;
TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Mus sp. (mouse)
PN JP 2001069994-A/22
PD 21-MAR-2001
PF 29-JUN-2000 JP 2000195911
PR
PI YASUAKI ITO, KAZUNORI NISHI, KAZUHIRO OGI, SHOICHI OKUBO, PI
SHINICHI MOGI,
PI YUKO NOGUCHI, KOJI YOSHIMURA, HIDEYUKI TANAKA
PC C12N15/09, A61K38/00, A61K45/00, A61K48/00, A61P9/00, A61P19/02, PC
A61P19/08,
PC C07K14/47, C07K16/18, C12N1/21, C12N5/10, G01N33/15, G01N33/50, PC
G01N33/53//
PC C12P21/08, C12N15/00, A61K37/02, C12N5/00
CC
FH Key Location/Qualifiers
FT source 1..947
FT /organism="Mus sp. (mouse)"
FT /location="Qualifiers"
FEATURES
Source Location/Qualifiers
1..947
/organism="Mus sp."
/db_xref="taxon:10095"
BASE COUNT 279 a 158 c 221 g 289 t
ORIGIN

Query Match 76.4%; Score 252.2; DB 6; Length 947;
Best Local Similarity 85.4%; Pred. No. 7.6e-65;
Matches 281; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 1 catggaataattatgacgcgtctagcttccaagaagctctgtgcagatgatgagtgctc 60
Db 65 CATGGTGATTTTATGGATAAACTTCTTCTTAAGAAGTTGTGTGCGGATGAGGAGTGTGTC 124
Qy 61 tatactattctctgctagtgctcaagaattataatgcccgagctgtagattcatt 120
Db 125 TATACTATTCTCTGGCAAGACACAGAGATTACATTACATGCCACAGACTGTAGGTTTCATC 184
Qy 121 aacgttaaaaggcagcagatctatgtgtactcaaaagctgtgtaaaagaaaatggagct 180
Db 185 GATGTCACAGAAAGGCGACGAGATCTATGTTTACTCAAGCTGGTTAACAGAAAACGAGCT 244
Qy 181 ggagaattttgggctggcagtgcttctatggtgatggccaggacagatgggagtcgtgggt 240
Db 245 GGAGAGTTTTGGCGTGGCAGTCTTTATGCTGACCCAGGATGAGATGGGAATTGTAGCT 304
```


Db 132 TATACATATTTCTGCGAAGAGACACAGGAGATTACAATGCCCGACACTGTAGGTTTCATC 191
QY 121 aacgttaaaagggcagcagatctatgtactcaaaagctgtaaaagaaatggagct 180
Db 192 GATGTCAGAAAGGGCAGCAGATCTATGTTACTCCAGCTGGTAACAGAAACGGAGCT 251
QY 181 ggaagaattttggcgtggcagtggtttatgtgtatggccagagcagagatgggagtcgtgggt 240
Db 252 GGAGAGTTTGGCGTGGCAGTGTATGTTGACACACAGGATGAGATGGGAATTGAGGT 311
QY 241 tatttccccaggaactgttcaaggaacacagcgtgtgtatccaggaagactccaaggaagtt 300
Db 312 TATTTCCCGCAACCTTGGTGAAGAGCAGCGGTATACAGGAGGCCACCAAGGAGATC 371
QY 301 ccaccacgggatattgactctcttcgcga 329
Db 372 CCAACCCAGGATATTGACTTCTTCGTGA 400

RESULT 13

BD010836
LOCUS BD010836 330 bp DNA linear PAT 31-JAN-2002
DEFINITION Novel polypeptide and DNA thereof.
ACCESSION BD010836
VERSION BD010836.1 GI:18639209
KEYWORDS JP 2001069994-A/37.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 330)
AUTHORS Ito,Y., Nishi,K., Ogi,K., Okubo,S., Mogi,S., Noguchi,Y.,
Yoshimura,K. and Tanaka,H.
TITLE Novel polypeptide and DNA thereof
JOURNAL Patent: JP 2001069994-A 37 21-MAR-2001;
TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Rattus sp. (rat)
PN JP 2001069994-A/37
PD 21-MAR-2001
PF 29-JUN-2000 JP 2000195911
PI YASUAKI ITO,KAZUNORI NISHI,KAZUHIRO OGI,SHOICHI OKUBO, PI
SHINICHI MOGI,
PI YUKO NOGUCHI,KOJI YOSHIMURA,HIDEYUKI TANAKA
PC C12N15/09,A61K38/00,A61K45/00,A61K48/00,A61P9/00,A61P19/02, PC
A61P19/08,
PC C07K14/47,C07K16/18,C12N1/21,C12N5/10,G01N33/15,G01N33/50, PC
G01N33/53//
PC C12P21/08,C12N15/00,A61K37/02,C12N5/00
CC
FH key Location/Qualifiers
FT source 1..330
FT /organism='Rattus sp. (rat)'.
FEATURES
source
1..330
/organism='Rattus sp.'
/db_xref='taxon:10118'
BASE COUNT 91 a 62 c 91 g 86 t
ORIGIN

Query Match 75.5%; Score 249; DB 6; Length 330;
Best Local Similarity 84.8%; Pred. No. 6e-64;
Matches 279; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 catggaattttatgacgcgttagcttccaagaagactctgtgcagatgatagtgctc 60
Db 1 CATGGCATGTTTATGGATAAACTTTCTTCTTAAGAAAGTTGTGTGCAGATGAGGAGTGTGC 60
QY 61 tatactatttcttggtctgagtcgaagattataatgcccgagactgtgattcatt 120
Db 61 TATACCATTTCTCTGCGAAGAGCAGCAGGAGACTACATATGCCCGGACTGTAGGTTTCATC 120

QY 121 aacgttaaaagggcagcagatctatgtactcaaaagctgtaaaagaaatggagct 180
Db 121 ATGTGTCAGAAAGGGCAGCAGATCTATGTTTATTCAGCTGGTAACAGAAATGGAGCT 180
QY 181 ggaagaattttggcgtggcagtggtttatgtgtatggccagagcagagatgggagtcgtgggt 240
Db 181 GGGGCATTTGGCGTGGCAGTGTATGTTGACACACAGGATGAGATGGGAATTGTTGGGT 240
QY 241 tatttccccaggaactgttcaaggaacacagcgtgtgtatccaggaagactccaaggaagtt 300
Db 241 TATTTCCCGCAACCTTGGTGAAGAGCAGCGGTATACAGGAGGCCACCAAGGAGATC 300
QY 301 ccaccacgggatattgactctcttcgcga 329
Db 301 CCAACCCAGGATATTGACTTCTTCGTGA 329

RESULT 14

BD010835
LOCUS BD010835 384 bp DNA linear PAT 31-JAN-2002
DEFINITION Novel polypeptide and DNA thereof.
ACCESSION BD010835
VERSION BD010835.1 GI:18639208
KEYWORDS JP 2001069994-A/36.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 384)
AUTHORS Ito,Y., Nishi,K., Ogi,K., Okubo,S., Mogi,S., Noguchi,Y.,
Yoshimura,K. and Tanaka,H.
TITLE Novel polypeptide and DNA thereof
JOURNAL Patent: JP 2001069994-A 36 21-MAR-2001;
TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Rattus sp. (rat)
PN JP 2001069994-A/36
PD 21-MAR-2001
PF 29-JUN-2000 JP 2000195911
PI YASUAKI ITO,KAZUNORI NISHI,KAZUHIRO OGI,SHOICHI OKUBO, PI
SHINICHI MOGI,
PI YUKO NOGUCHI,KOJI YOSHIMURA,HIDEYUKI TANAKA
PC C12N15/09,A61K38/00,A61K45/00,A61K48/00,A61P9/00,A61P19/02, PC
A61P19/08,
PC C07K14/47,C07K16/18,C12N1/21,C12N5/10,G01N33/15,G01N33/50, PC
G01N33/53//
PC C12P21/08,C12N15/00,A61K37/02,C12N5/00
CC
FH key Location/Qualifiers
FT source 1..384
FT /organism='Rattus sp. (rat)'.
FEATURES
source
1..384
/organism='Rattus sp.'
/db_xref='taxon:10118'
BASE COUNT 98 a 72 c 109 g 105 t
ORIGIN

Query Match 75.5%; Score 249; DB 6; Length 384;
Best Local Similarity 84.8%; Pred. No. 6.1e-64;
Matches 279; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 catggaattttatgacgcgttagcttccaagaagactctgtgcagatgatagtgctc 60
Db 55 CATGGCATGTTTATGGATAAACTTTCTTCTTAAGAAAGTTGTGTGCAGATGAGGAGTGTGC 114
QY 61 tatactatttcttggtctgagtcgaagattataatgcccgagactgtgattcatt 120
Db 115 TATACCATTTCTCTGCGAAGAGCAGCAGGAGACTACATATGCCCGGACTGTAGGTTTCATC 174

QY 121 aacgttataaaagggcagcagatctatgttactcaagctgtgtaaaagaaatggagct 180
Db 175 AATGTCAGAAAGGGCAGCAGATCTATGTTTATTCGAAGCTGGTAACAGAAAATGGAGCT 234
QY 181 ggagaattttgggctggcagtggtttatggatggccagacagagatggagtcgtgggt 240
Db 235 GGGCATTCTGGGCTGGCAGTGTATGTTGACCAACAGGATGAGATGGGAATTGTGGGT 294
QY 241 tatttccccaggaacttgggtcaaggaacagcgtgtgtaccaggaagctaccaaggaagt 300
Db 295 TATTTCCCAAGCAACTTGGTTAGAGAGCAACGAGTGTACCAAGAGGGCCACCAAGGAGATT 354
QY 301 ccaacacagatattgacttcttctgcga 329
Db 355 CCAACACGGGATATTGACTTCTTCTGTGA 383

RESULT 15

BD010830
CUS BD010830 307 bp DNA linear PAT 31-JAN-2002
INITIATION Novel polypeptide and DNA thereof.
ACCESSION BD010830
VERSION BD010830.1 GI:18639203
KEYWORDS JP 2001069994-A/31.
SOURCE Rattus sp.
ORGANISM Rattus sp.

REFERENCE 1 (bases 1 to 307)
AUTHORS Ito,Y., Nishi,K., Ogi,K., Okubo,S., Mogi,S., Noguchi,Y.,
Yoshimura,K. and Tanaka,H.
TITLE Novel polypeptide and DNA thereof
JOURNAL Patent: JP 2001069994-A 31 21-MAR-2001;
TAKEDA CHEMICAL INDUSTRIES LTD

COMMENT OS Rattus sp. (rat)
PN JP 2001069994-A/31
PD 21-MAR-2001
PF 29-JUN-2000 JP 2000195911

PI YASUAKI ITO,KAZUNORI NISHI,KAZUHIRO OGI,SHOICHI OKUBO, PI
SHINICHI MOGI,
PI YUKO NOGUCHI,KOJI YOSHIMURA,HIDEYUKI TANAKA
PC C12N15/09,A61K38/00,A61K45/00,A61K48/00,A61P9/00,A61P19/02, PC
A61P19/08,
PC C07K14/47,C07K16/18,C12N1/21,C12N5/10,G01N33/15,G01N33/50, PC
G01N33/53//
PC C12P21/08,C12N15/00,A61K37/02,C12N5/00

CC Key Location/Qualifiers
FH Source 1..307
FT /Organism='Rattus sp. (rat)',
Location/Qualifiers

FEATURES

source
1..307
/organism='Rattus sp.'
/db_xref='taxon:10118'
84 a 56 c 90 g 77 t

BASE COUNT
ORIGIN

Query Match 65.0%; Score 214.6; DB 6; Length 307;
Best Local Similarity 82.1%; Pred. No. 1.2e-53;
Matches 247; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 1 catggaattatttgaccgtctagcttccaaagctctgtgcagatgatgagtggtc 60
Db 7 CATGGTGTATTTATGGATAAACTTCTTCTTAAGAAAGTTGTGTCAGATGAGGAGTGTGTC 66
QY 61 tatactatttctctggttagtgctcaagaagattataatccccggactgtagattcatt 120
Db 67 TATACCAATTTCTCTGGCAAGACGACAGAAAGACTACAAATGCCCGGACTGTAGTTTCATC 126
QY 121 aacgttataaaagggcagcagatctatgttactcaagctgtgtaaaagaaatggagct 180

Db 127 AATGTCAGAAAGGGCAGCAGATCTATGTTTATTCGAAGCTGGTAACAGAAAATGGAGCT 186
QY 181 ggagaattttgggctggcagtggtttatggatggccagacagagatggagtcgtgggt 240
Db 187 GGGCATTCTGGGCTGGCAGTGTATGTTGACCAACAGGATGAGATGGGAATTGTGGGT 246
QY 241 tatttccccaggaacttgggtcaaggaacagcgtgtgtaccaggaagctaccaaggaagt 300
Db 247 TATTTCCCAAGCAACTTGGTTAGAGAGCAACGAGTGTACCAAGAGGGCCACCAAGGAGAT 306
QY 301 c 301
Db 307 C 307

Search completed: September 23, 2002, 14:58:47
Job time: 2061 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2002, 13:38:35 ; Search time 1903.73 Seconds
(without alignments)
2339.615 Million cell updates/sec

Title: US-10-019-455a-23
Perfect score: 330
Sequence: 1 catggaattatttgaccg.....atattgactttcttcgcgag 330

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

al number of hits satisfying chosen parameters: 27472414

Minimum db seq length: 0
Maximum db seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	252.2	76.4	696	9 BB611549	BB611549 BB611549
2	190.4	57.7	527	9 BE236443	BE236443 144645 MA
3	157.4	47.7	349	9 AW023324	AW023324 df53d01.y
4	131.6	39.9	338	9 AW021345	AW021345 df21c01.y
5	110	33.3	678	9 BB647928	BB647928 BB647928
6	107.2	32.5	492	9 AA22297	AA22297 mv77b12.r
7	82	24.8	576	10 B1492218	B1492218 df21c01.w
8	64.6	19.6	533	10 W74647	W74647 zd77e05.s1
9	63.4	19.2	373	10 BF691795	BF691795 602248344
10	63.4	19.2	890	10 BG766328	BG766328 602739014
11	63.4	19.2	1140	10 BG765502	BG765502 602739414
12	62.4	18.9	657	12 AG141467	AG141467 Pan trogl
13	61.8	18.7	430	10 BF439750	BF439750 nad13c10.
14	61.8	18.7	442	9 AA282143	AA282143 zt02b05.s
15	61.6	18.7	388	10 BE525492	BE525492 uul19d01.y
16	61.6	18.7	440	9 BB749350	BB749350 BB749350
17	61.6	18.7	464	9 AI323423	AI323423 mj58f08.x

18	61.6	18.7	465	9 AA003262	AA003262
19	61.6	18.7	555	10 B1851688	B1851688 603377922
20	61.4	18.6	381	9 AI620915	AI620915 tu05e09.x
21	61.4	18.6	698	10 BE960967	BE960967 601648520
22	61.2	18.5	452	9 AA428926	AA428926 zv77d08.r
23	60.2	18.2	453	9 AA627297	AA627297 nq68g01.s
24	60.2	18.2	457	10 BE864464	BE864464 UI-M-BH1-
25	60	18.2	409	9 AW046043	AW046043 UI-M-BH1-
26	58.8	17.8	496	9 AW444509	AW444509 UI-H-B13-
27	57.4	17.4	425	10 W75984	W75984 zd58e03.r1
28	56.6	17.2	300	9 AU099830	AU099830 AU099830
29	56.4	17.1	396	10 W80195	W80195 me87d07.r1
30	56.4	17.1	475	9 AA982842	AA982842 ub59e03.r
31	55.4	16.8	391	9 AI275598	AI275598 qm48e04.x
32	55.2	16.7	481	9 AI323068	AI323068 mj58f08.y
33	54.2	16.4	437	9 AV592759	AV592759 AV592759
34	54.2	16.4	517	10 BE665724	BE665724 154861 MA
35	53.4	16.2	392	10 W54756	W54756 md10g04.r1
36	53.4	16.2	750	10 BF160660	BF160660 601768554
37	53.2	16.1	448	9 AA537462	AA537462 vk47d10.r
38	53.2	16.1	767	10 BF977676	BF977676 602147484
39	52.8	16.0	345	9 AW326170	AW326170 18352 MAR
40	52.6	15.9	589	9 AV169321	AV169321 AV169321
41	52.2	15.8	452	9 AA960553	AA960553 ub59e03.s
42	52.2	15.8	668	10 BF977955	BF977955 602148428
43	52	15.5	386	10 W54019	W54019 md09a01.r1
44	51.2	15.5	328	9 AA058045	AA058045 mj58f08.r
45	50.2	15.2			

ALIGNMENTS

RESULT 1

BB611549
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

BB611549 RIKEN full-length enriched, 13 days embryo head Mus
musculus cDNA clone 3110083012 5', mRNA sequence.

BB611549 GI:15393547

EST.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 696)

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.

, Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda

, M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,

Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki

, D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,

Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,

Muramatsu,M. and Hayashizaki,Y.

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gscc.riken.go.jp,

URL:http://genome.gsc.riken.go.jp/

Carninci,P., Shibata,Y., Hayatsu,M., Muramatsu,M. and Hayashizaki,Y.

, M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

waqi,K., Fujikawa,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura

, S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and

Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000).

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T., Ishii,Y. and Hayashizaki,Y.

Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Func. Genomics 2 pre, L72-L86 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

1..696

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone_lib="J110083012"

/clone_lib="RIKEN full-length enriched, 13 days embryo head"

/tissue_type="head"

/dev_stage="13 days embryo"

/lab_host="SOLR"

/note="Site 1: XhoI; Site 2: SstI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5',

GAGAGAGAGAGATCCCAAGAGCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 50.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGATGATTAATTAATCCCGCCCCCCC 3']. cDNA was cleaved with XhoI and SstI."

BASE COUNT 204 a 122 c 176 g 194 t

ORIGIN

Query Match 76.4%; Score 252.2; DB 9; Length 696;

Best Local Similarity 85.4%; Pred. No. 3.1e-62;

Matches 281; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

1 catggaattttatgacccgtctagcttccaagaagctctgtgcagatgatgagtgtgc 60

||||| ||||||| || ||||| ||||||| ||||||| ||||||| |||||||

Db 78 CATGGTGATATTTATGATAAACTTCTTCTAAGAGATTGTGTGCGGATGAGGAGTGTGTC 137

QY 61 tatactatttctctgctagctgaagattataatgccccggaactgtagattcatt 120

||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 138 TATACATTTTCTCTGGCAGAGACAGGAGATTAACAATGCCCGACACTGTAGGTTTCATC 197

QY 121 aacgttaaaaaagggcagcagatctatgtgtactcaaaagctgtgtaaaagaaatggagct 180

||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 198 GATGTCAAGAAAGGGCAGCAGATCTATGTTACTCCAGCTGGTACAGAAAACGGAGCT 257

QY 181 ggagaattttggcctggcagtgcttatttgatgagcaggaagatgggagtcgtggtg 240

||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 258 GGAGAGATTTTGGCTGGCAGTGTGTTATGTGTACACAGGATGAGATGGGAATTGTAGCT 317

QY 241 tatttccccagaactgtctcaaggaacagcggtgtgtaccaggaagctcccaagaagtt 300

||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 318 TATTTCCCGAGCAACTTGGTGAGGAGACGCGGTGTATACAGAGAGGCCACCAAGGAGATC 377

QY 301 cccaccacggatattgaactctcttcgcga 329

||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 378 CCAACACCGGATATTGACTTCTCTCTGA 406

RESULT 2

BE236443

LOCUS

DEFINITION

144845 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION

BE236443

VERSION

BE236443.1 GI:9021161

KEYWORDS

EST.

SOURCE

COW.

ORGANISM

Bos taurus

Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

1 (bases 1 to 527)

Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

21180013

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCAGTCACGACG

Plate: 54 row: C column: 18

Seq primer: AFTTAGGTGACACTATAG.

FEATURES

source

1..527

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_lib="MARC 4BOV"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site 1: XbaI; Site 2: XhoI; Library made from pooled tissue from day 20 and day 40 embryos."

BASE COUNT 173 a 94 c 126 g 134 t

ORIGIN

Query Match 57.7%; Score 190.4; DB 9; Length 527;

Best Local Similarity 78.2%; Pred. No. 2e-44;

Matches 258; Conservative 0; Mismatches 36; Indels 36; Gaps 1;

QY 1 catggaattttatgacccgtctagcttccaagaagctctgtgcagatgatgagtgtgc 60

||||| ||||||| || ||||| ||||||| ||||||| ||||||| |||||||

Db 89 CATGGAATATTTATGACAGACTGTGTTCCAAGAAGCTGTGCAGATGATGAATGTGC 148

QY 61 tatactatttctctgctagctgaagattataatgccccggaactgtagattcatt 120

||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 149 TATACATTTTCTCTGCGCAGAGCTCAAGAAGATTACAATGCTCCCGACGTAGATTTCAT 208

QY 121 aacgttaaaaaagggcagcagatctatgtgtactcaaaagctgtgtaaaagaaatggagct 180

||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 209 ACAGTTAAAAAGGACAGTGGATCTATGTTTACTCAAGCTGGTCTATGCATCATCTCT 268

QY 181 ggagaattttggcctggcagtgcttatttgatggcaggaagatgggagtcgtggtg 240

||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 269 G-----AGGATGAAATGGGAACCGTGGGT 292

QY 241 tatttccccagaacttggtcaaggaacagcggtgtgtaccaggaagctcccaaggagtt 300

||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||


```

/clone_lib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT      219 a      244 c      291 g      136 t
ORIGIN

Query Match      19.2%; Score 63.4; DB 10; Length 890;
Best Local Similarity 58.2%; Pred. No. 1.2e-07;
Matches 153; Conservative 0; Mismatches 101; Indels 9; Gaps 2;

12 tatggaccgtctagctccaagaagctctgtgcagatgatgagtggtctatactatttc 71
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 101 TATGCCCAAGCTGGCTGACCGGAAGCTGTGTGGGACCGAGAGTGCAGCCACCCTATCTC 160

QY 72 tctggctagtctcaagaattataatgcccgagactgtagattcatttaacgttaaaaa 131
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 161 CATGGCTGTGGCCCTTCAGGACTACATGCCCCCGAGCTGCCGATTCCTGACCATTACCG 220

QY 132 agggcagcagatctatgttactcacaagctggttaaaagaaatggagctggagaatttg 191
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 221 GGCCCAAGTGGTGTATGCTCTCTCCAAAGCTG-----AAGGGCCGTGGGGCGCTCTCTG 274

QY 192 ggtcggcagtggttattggtgagtcggcagacgagatgggag-----tcgtgggttatttc 248
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 275 GGGAGGCAGCGTTTCAGGGAGATTACTATGGAGATCTGCTGCTGCGCTGGGCTATTTC 334

QY 249 caggaacttggtcaagaaacagc 271
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 335 CAGTAGATTGTCGAGAGGACC 357

RESULT 11
BG765502
LOCUS      602739414F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4869267 5',
DEFINITION mRNA sequence.
ACCESSION BG765502
VERSION   BG765502.1 GI:14076155
KEYWORDS EST.
SOURCE    human.
ORGANISM  Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE     NIH-MGC http://mgc.nci.nih.gov/
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgapbs@remail.nih.gov
          Tissue Procurement: ATCC/DCTD/DTF
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: L10M1740 row: 1 column: 04
          High quality sequence stop: 509.
          Location/Qualifiers
            1..1140
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:4869267"
FEATURES
source

```

```

/clone_lib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT      251 a      336 c      324 g      228 t
ORIGIN

Query Match      19.2%; Score 63.4; DB 10; Length 1140;
Best Local Similarity 58.2%; Pred. No. 1.4e-07;
Matches 153; Conservative 0; Mismatches 101; Indels 9; Gaps 2;

12 tatggaccgtctagctccaagaagctctgtgcagatgatgagtggtctatactatttc 71
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 101 TATGCCCAAGCTGGCTGACCGGAAGCTGTGTGGGACCGAGAGTGCAGCCACCCTATCTC 160

QY 72 tctggctagtctcaagaattataatgcccgagactgtagattcatttaacgttaaaaa 131
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 161 CATGGCTGTGGCCCTTCAGGACTACATGCCCCCGAGCTGCCGATTCCTGACCATTACCG 220

QY 132 agggcagcagatctatgttactcacaagctggttaaaagaaatggagctggagaatttg 191
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 221 GGCCCAAGTGGTGTATGCTCTCTCCAAAGCTG-----AAGGGCCGTGGGGCGCTCTCTG 274

QY 192 ggtcggcagtggttattggtgagtcggcagacgagatgggag-----tcgtgggttatttc 248
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 275 GGGAGGCAGCGTTTCAGGGAGATTACTATGGAGATCTGCTGCTGCGCTGGGCTATTTC 334

QY 249 caggaacttggtcaagaaacagc 271
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 335 CAGTAGATTGTCGAGAGGACC 357

RESULT 12
AG141467/c
LOCUS      AG141467
DEFINITION Pan troglodytes DNA, clone: RP43-001J02.T7, genomic survey
sequence.
ACCESSION AG141467
VERSION   AG141467.1 GI:16671145
KEYWORDS GSS: GSS (genome survey sequence).
SOURCE    Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee
          Male BAC Library clone:RP43-001J02.T7.
ORGANISM  Pan troglodytes
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Euthera; Primates; Catarrhini; Hominidae; Pan.
TITLE     Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
JOURNAL   Totoki, Y., Watanabe, H. and Sakaki, Y.
COMMENT   BAC end sequences of Library RPCI-43
          Unpublished
          2 (bases 1 to 657)
          Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
          Totoki, Y., Watanabe, H. and Sakaki, Y.
          Direct Submission
          Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
          and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
          1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
          (E-mail:chimpbes@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
          Tel:81-45-503-9111, Fax:81-45-503-9170)
          Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
          end was generated during the R&D process and may have higher chance
          of clone tracking errors.
          PRIMERS
            Sequencing: T7

```

```
LIBRARY
Vector      : pBac3.6
R.Site 1    : EcoRI
R.Site 2    : EcoRI
Location/Qualifiers
1..657
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-001J02.T7"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPC1-43 Chimpanzee Male BAC Library"
201 a 143 c 117 g 189 t 7 others

BASE COUNT      201 a 143 c 117 g 189 t 7 others
ORIGIN
|||||
Query Match      18.9%; Score 62.4; DB 12; Length 657;
Best Local Similarity 98.4%; Pred. No. 2.1e-07;
Matches 63; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 catggaattattatgacacctagcttccaaagaagctctgtgcagatgatgagtgtgc 60
|||||
248 CATGGAATATTTATGACCGCTAGCTTCCAGAGACCTCTGTGCAGATGATGAGTGTGC 189
Db
|||||
QY 61 tata 64
|||
Db 188 TGTA 185

RESULT 13
BF439750/c
LOCUS
DEFINITION
BF439750
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 430)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco.
Location/Qualifiers
1..430
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3365371"
/clone_lib="NCI-CGAP_Lu24"
/tissue_type="carcinoid"
/lab_host="DH10B"
/notes="Organ: lung; Vector: pTT73D-Pac (Pharmacia) with a modified
polylinker; Plasmid DNA from the normalized
library NCI-CGAP_Lu5 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (cloneIDs

1414920-1417991 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo.
BASE COUNT      100 a 124 c 125 g 81 t
ORIGIN
|||||
Query Match      18.7%; Score 61.8; DB 10; Length 430;
Best Local Similarity 57.8%; Pred. No. 2.7e-07;
Matches 152; Conservative 0; Mismatches 102; Indels 9; Gaps 2;

QY 12 tatgaccgctctagcttccaaagaactctgtgcagatgatgagtgtgtctatactatttc 71
|||||
388 TATGCCCAAGCTGGGTGACCGGAAGCTGTGTCCGACACAGGAGTGCAGCCACCTATCTC 329
Db
|||||
QY 72 tctgctagtgctcaagaagattataatgccccgcgactgtagattcattacaagctlaaaaa 131
|||||
328 CATGCTGTGGCCCTTCAGGACTACATGTCGCCGACTGCCGATTCCTGACCATTCACCG 269
Db
|||||
QY 132 agggcagcagatctatgtgtactcaaaagctgtgtataaagaataatggagctggagaattttg 191
|||||
268 GGGCCCAAGTGGTGTATGCTTCTTCCAAAGCTG-----AAGGGCCGTGGCGCGCTCTTCG 215
Db
|||||
QY 192 ggcctggcagtgcttctgtgtatggtgcccagagacagatggag---tcgtgggtatttccc 248
|||||
214 GGGAGGCGACGCTTCAGGGAGATTACTATGAGATCTGGCTGCTCGCTGGGCTATTTCCC 155
Db
|||||
QY 249 caggaaacttgctcaaggaaacgc 271
|||||
Db 154 CAGTAGCATTTGTCGAGAGGACC 132

RESULT 14
AA282143/c
LOCUS
DEFINITION
AA282143
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 442)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Insert Length: 1148 Std Error: 0.00
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 337.
Location/Qualifiers
1..442
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:711921"
/clone_lib="NCI-CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/notes="vector: pTT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not 1; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
15'-TGTACCAATCTGAAGTGGAGCGCGCCCTCATTTTTTTTTTTTTTTT-3'
```

1. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library is normalized. Library and Eco RI sites of the modified p7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
ORIGIN

92 a 119 c 121 g 110 t

Query Match

Best Local Similarity 18.7%; Score 61.8; DB 9; Length 442;

Matches 152; Conservative 0; Mismatches 102; Indels 9; Gaps 2;

QY 12 tatggaccgtctagcttccaagaagctctgtgcagatgatgagtggtctatactatttc 71

Db 416 TATGCCCAAGCTGGCTGACCGAAGCTGTGTGCGGACCAAGAGTGCGAGCCACCCCTATCTC 357

QY 72 tctgggtagtgctcaagaagattataatccccggagcttagattcattaaagttaaaaa 131

356 ACATGCTGTGGCCCTTCAGGACTACATGCCCCCGACTCCGATTCCTGACCATTCACCG 297

132 agggcgagcatctatgttactcaagctgtgtaaaagaaatggagctggagaattttg 191

Db 296 GGGCCCAAGTGGTGTATGTCTTCTCCAAAGTG-----AACGGCCGTGGCGGCTCTCTG 243

QY 192 gctggcagctgtttatgttgatggccagacagagatggag---tcgtgggttatttccc 248

Db 242 GGGAGCAGCGTTTCAGGAGATTACTATGAGATCTGGCTGCTCGCTGGGCTATTTTCCC 183

QY 249 caggaacttggtcaagaagacagc 271

Db 182 CAGTAGCATTTGTCGAGAGACC 160

RESULT 15

BE625492

LOCUS

uul9d01.y1 Soares mammary_gland_NLMG Mus musculus cDNA clone
IMAGE:3372385 5' similar to SW:MIA_MOUSE Q61865 MELANOMA DERIVED
GROWTH REGULATORY PROTEIN PRECURSOR ;, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:1081989

Seq primer: -40RP from Gibco

High quality sequence stop: 382.

Location/Qualifiers

1..388

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone_lib="Soares_mammary_gland_NLMG"

/sex="female (lactating)"

/tissue_type="mammary gland"

/lab_host="DH10B"

/note="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not

I and cloned into the Not I and Eco RI sites of the
modified p7T3 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
ORIGIN

80 a 105 c 110 g 93 t

Query Match

Best Local Similarity 18.7%; Score 61.6; DB 10; Length 388;

Matches 174; Conservative 0; Mismatches 139; Indels 9; Gaps 2;

QY 12 tatggaccgtctagcttccaagaagctctgtgcagatgatgagtggtctatactatttc 71

Db 6 TATGCCCAAGCTGGCTGACCTGGAAGCTGTGTGCGGACGAGGAATGCGAGCCATCTTATCTC 65

QY 72 tctgggtagtgctcaagaagattataatccccggagcttagattcattaaagttaaaaa 131

Db 66 CATGGCTGTGGCCCTCCAGGACTACGTGGCCCTGATTGCCGCTTCTTGACTATATATAG 125

QY 132 agggcgagcatctatgttactcaagctgtgtaaaagaaatggagctggagaattttg 191

Db 126 GGGCCCAAGTGGTGTATGTCTTCTCCAAAGTG-----AAGGGCCGTGGCGGCTCTTCTG 179

QY 192 gctggcagctgtttatgttgatggccagacagagatggag---tcgtgggttatttccc 248

Db 180 GGGAGCAGCTGTTTCAGGAGATTACTATGAGACCTGGAGCCCGCTGGGCTATTTTCCC 239

QY 249 caggaacttggtcaagaagacagcgtgtgtaccagggaagctaccaaggaagtccaccac 308

Db 240 CAGTAGCATTTGTCGAGAGACCCTGACTCTGAACACCTGGCAAAATTTGATATGAAGACCGA 299

QY 309 ggatattgacttcttctgcgag 330

Db 300 TCAATGGGATTTTCTACTGCCAG 321

Search completed: September 23, 2002, 14:24:13

Job time: 2738 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2002, 15:36:31 ; Search time 224.07 Seconds
(without alignments)
2528.593 Million cell updates/sec

Title: us-10-019-455a-23

Perfect score: 330

Sequence: 1 catgaatatttgaccg.....atattgactctctgcgag 330

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

al number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- N_Geneseq_032802.*
- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
 - 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
 - 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
 - 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
 - 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
 - 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
 - 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
 - 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
 - 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
 - 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
 - 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
 - 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
 - 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
 - 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
 - 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
 - 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
 - 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
 - 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
 - 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
 - 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
 - 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
 - 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
 - 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
 - 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	330	100.0	330	22	AAF59079
2	330	100.0	384	22	AAF59065
3	330	100.0	387	24	AA517583
4	330	100.0	426	22	AAH26341
5	330	100.0	891	22	AAH98228
6	330	100.0	891	22	AAH26342
7	330	100.0	923	22	AAF59083
8	330	100.0	1201	22	AAH26343
9	252.2	76.4	330	22	AAF59080

10	252.2	76.4	384	22	AAF59088	Mouse MLP nucleoti
11	252.2	76.4	947	22	AAF59084	Mouse MLP nucleoti
12	249	75.5	330	22	AAF59099	Rat MLP nucleotide
13	249	75.5	384	22	AAF59098	Rat MLP nucleotide
14	214.6	65.0	307	22	AAF59093	Rat MLP nucleotide
15	193.8	58.7	261	22	AAF59092	Rat MLP nucleotide
16	63.4	19.2	433	22	AAH47783	Recombinant human
17	63.4	19.2	459	16	AAQ84050	Sequence encoding
18	63.4	19.2	459	22	AAI70083	Melanoma inhibitor
19	63.4	19.2	459	22	AAI18732	Human antisenase ol
20	63.2	19.2	581	16	AAQ84052	Sequence encoding
21	57.4	17.4	330	16	AAQ84061	Sequence encoding
22	54.2	16.4	305	16	AAQ84055	Amplified fragment
23	53.6	16.2	300	20	AAI14828	Human gene express
24	53.6	16.2	429	22	AAI22695	Human CDNA encodin
25	53.6	16.2	884	22	AAI22459	Human CDNA encodin
26	53.6	16.2	1263	21	AAI21245	Human TANGO 130 cd
27	52	15.8	417	22	AAH99775	Human protein enco
28	46	13.9	1060	22	AAH92140	Human PRQ19670 CDN
29	46	13.9	1061	22	AAH46205	Human DNA encoding
30	46	13.9	2886	21	AAI21244	Murine TANGO 130 c
31	44.2	13.4	421	21	AAI21244	Human secreted pro
32	44.2	13.4	2037	22	AAI21244	Human DNA encoding
33	40.4	12.2	596	16	AAQ84066	Sequence encoding
34	40.4	12.2	3565	16	AAQ84051	Sequence encoding
35	35.6	10.8	13732	24	ABL33820	Human immune syste
36	35	10.6	582	20	AAV86447	EST clone AW728.
37	35	10.6	4292	20	AAH80481	Human secreted pro
38	35	10.6	4292	22	AAH80481	Human CDNA encodin
39	35	10.6	4292	24	ABA90896	Human polynucleoti
40	34.4	10.4	2003	23	ABL08217	Drosophila melanog
41	34.4	10.4	6169	23	ABL08216	DNA encoding SNV-e
42	33.8	10.2	1030	22	AAH61509	Recombinant fructo
43	33.6	10.2	1136	17	AAI35254	Tomato LesNF1 (suc
44	33.6	10.2	1929	22	AAH62011	Endotoxin CryV gen
45	33.6	10.2	2165	18	AAH59702	

ALIGNMENTS

RESULT 1

AAF59079
ID AAF59079 standard; DNA; 330 BP.

AC AAF59079;

XX 23-APR-2001 (first entry)

DT Human MLP nucleotide sequence SEQ ID NO:23.

DE MLP; MIA; melanoma inhibitory activity; cancer; bone disease;

KW joint disease; pathologic angiogenesis; diagnosis; antiinflammatory;

KW cardiant; gene therapy; secretory cell function regulator; promoter;

KW inhibitor; ds.

XX Homo sapiens.

OS Homo sapiens.

XX WO200102564-A1.

XX 11-JAN-2001.

XX 29-JUN-2000; 2000WO-JP04278.

XX 30-JUN-1999; 99JP-0186718.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Itoh Y, Nishi K, Ogi K, Ohkubo S, Mogi S, Noguchi Y, Yoshimura K;

XX Tanaka H;

XX WPI; 2001-159271/16.

XX P-PSDB; AAB69126.

XX Safe, low-toxicity secretory cell function-regulatory protein and
PT encoded DNA, applicable as drugs, in diagnosis and development of
PT promoters and inhibitors for preventing or treating e.g. bone and joint
PT diseases -
XX
XX
PS Claim 8; Page 97; 111pp; Japanese.
PS
XX
CC The present invention describes novel MLP proteins and their encoding
CC DNAs. The MLP proteins and DNAs have antiinflammatory and cardiant
CC activities, and can be used in gene therapy and as secretory cell
CC function regulators. The MLP proteins and DNAs can be used in drugs, in
CC the diagnosis and development of promoters and inhibitors for preventing
CC or treating bone and joint diseases as well as pathologic angiogenesis.
CC AAF59063 to AAF59099 and AAB69122 to AAB69132 represent sequences used
CC in the exemplification of the present invention.
XX
XX Sequence 330 BP; 91 A; 60 C; 91 G; 88 T; 0 other;
XX

Query Match 100.0%; Score 330; DB 22; Length 330;
Best Local Similarity 100.0%; Pred. No. 2.8e-97;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 1 catggaatatttatggacccttagcttccaagaagctctgtgcagatgatgagtgctc 60
|||||
DB 1 catggaatatttatggacccttagcttccaagaagctctgtgcagatgatgagtgctc 60
|||||

QY 61 tatactattctctggctagtcgtccaagaagattataatgccccgactgtagattcatt 120
|||||
DB 61 tatactattctctggctagtcgtccaagaagattataatgccccgactgtagattcatt 120
|||||

QY 121 aacgttaaaaaaggcagcagatctatgttactcaaacgtgttaaaagaaatggagct 180
|||||
DB 121 aacgttaaaaaaggcagcagatctatgttactcaaacgtgtgttaaaagaaatggagct 180
|||||

QY 181 ggagaattttggctggcagctgtttatggtgatgcccagacgagatgaggagtcgtgggt 240
|||||
DB 181 ggagaattttggctggcagctgtttatggtgatgcccagacgagatgaggagtcgtgggt 240
|||||

QY 241 tatttccccaggaacttggccaaggacacgctgtgtaccaggaagctaccagaagaagt 300
|||||
DB 241 tatttccccaggaacttggccaaggacacgctgtgtaccaggaagctaccagaagaagt 300
|||||

QY 301 cccaccacggatatgtgacttctctgcgag 330
|||||
DB 301 cccaccacggatatgtgacttctctgcgag 330
|||||

RESULT 2
AAF59065
ID AAF59065 standard; DNA; 384 BP.
XX AC AAF59065;
XX
XX
DT 23-APR-2001 (first entry)
XX
DE Human MLP nucleotide sequence SEQ ID NO:4.
XX
XX MLP; MIA; melanoma inhibitory activity; cancer; bone disease;
KW joint disease; pathologic angiogenesis; diagnosis; antiinflammatory;
KW cardiant; gene therapy; secretory cell function regulator; promoter;
KW inhibitor; ds.
XX
OS Homo sapiens.
XX
XX WQ200102564-A1.
XX
XX 11-JAN-2001.
XX
XX 29-JUN-2000; 2000WO-IP04278.
XX
XX 30-JUN-1999; 99JP-0186718.
XX
XX

```
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..387
XX FT /*tag= a
XX FT /product= "Human secreted protein"
XX
XX WO200179454-A1.
XX
XX PD 25-OCT-2001.
XX
XX PF 11-APR-2001; 2001WO-US11797.
XX
XX PR 13-APR-2000; 2000US-196603P.
XX PR 24-APR-2000; 2000US-199417P.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX PA (SMIK ) SMITHKLINE BEECHAM PLC.
XX
XX Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z;
XX
XX WPI; 2002-061975/08.
XX DR P-PSDB; AAU09871.
XX
XX New secreted proteins or polypeptides, useful for treating e.g. cancer,
XX PT autoimmune diseases, wound healing disorder, infections, haematopoietic
XX PT disorders, inflammatory disorders, infertility, cancer -
XX
XX PS Claim 2; Page 44; 92pp; English.
XX
XX The invention relates to an isolated novel secreted polypeptide (I) and
XX CC polynucleotide (II). (I) and (II) are useful for treating cancer,
XX CC autoimmune diseases, wound healing disorder, infections, haematopoietic
XX CC disorders, inflammatory disorders, infertility, neurological and
XX CC psychiatric diseases, cardiovascular diseases, respiratory diseases,
XX CC renal diseases, or gastrointestinal diseases. These may also be used to
XX CC treat diseases, abnormalities and disorders caused by abnormal
XX CC expression, production, function and/or metabolism of the genes, as
XX CC vaccines for inducing immunological response in a mammal, and in
XX CC screening methods for detecting the effect of added compounds on the
XX CC production of mRNA and polypeptide in cells. The polypeptides can be used
XX CC as immunogens to produce antibodies immunospecific for the polypeptides,
XX CC and to identify membrane-bound or soluble receptors. The polynucleotides
XX CC may be used as diagnostic reagents, in chromosome localisation studies,
XX CC and in tissue expression studies. The present sequence represents the
XX CC coding sequence of novel human secreted protein #12.
XX
XX Sequence 387 BP; 101 A; 70 C; 106 G; 110 T; 0 other;
XX
XX Query Match 100.0%; Score 330; DB 24; Length 387;
XX Best Local Similarity 100.0%; Pred. No. 3.le-97;
XX Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 catggaattattgacgctgactctcaccagaagctctgtgcagatgatgagtgctc 60
XX |||||
XX DB 55 catggaattattgacgctgactctcaccagaagctctgtgcagatgatgagtgctc 114
XX |||||
XX QY 61 tatactattctctgctagtgctcaagaagattataatgcccgactgtagattcatt 120
XX |||||
XX DB 115 tatactattctctgctagtgctcaagaagattataatgcccgactgtagattcatt 174
XX |||||
XX QY 121 aacgttaaaaggcagcagatctatgtgtactcaaaagctgtgtaaaagaatggagct 180
XX |||||
XX DB 175 aacgttaaaaggcagcagatctatgtgtactcaaaagctgtgtaaaagaatggagct 234
XX |||||
XX QY 181 ggagattttggctggcagtgcttattatggtatgcccaggacagatgggagctcgtgggt 240
XX |||||
XX DB 235 ggagattttggctggcagtgcttattatggtatgcccaggacagatgggagctcgtgggt 294
XX |||||
XX QY 241 tatttccccagaaacttgtaagaaacagctgtgtaccaggaagctaccagaagatt 300
XX |||||
```

```
Db 295 tatttccccagaaacttgtaagaaacagctgtgtaccaggaagctaccagaagatt 354
QY 301 cccaccacggatattgacttcttctgtcgag 330
|||
Db 355 cccaccacggatattgacttcttctgtcgag 384
|||

RESULT 4
AAH26341
ID AAH26341 standard; cDNA; 426 BP.
XX
XX AAH26341;
XX
XX DT 02-OCT-2001 (first entry)
XX
XX DE Human growth regulatory-like polypeptide clone 16372272.
XX
XX KW Growth regulatory-like polypeptide; human; cartilage; melanoma;
XX KW neuroectodermal tumour; glioma; cancer; gene therapy; diagnosis;
XX KW ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200155332-A2.
XX
XX PD 02-AUG-2001.
XX
XX PF 25-JAN-2001; 2001WO-US02455.
XX
XX PR 25-JAN-2000; 2000US-0491404.
XX PR 02-MAY-2000; 2000US-0563786.
XX
XX (HYSE-) HYSEQ INC.
XX
XX MIze NK, Boyle BJ, Ford JE, Arterburn MC, Tang YT, Liu C;
XX Drmanac RT;
XX WPI; 2001-483233/52.
XX
XX Isolated human growth regulatory-like polypeptide useful for treating
XX e.g. Alzheimer's disease, cancer, autoimmune disorders,
XX hyperproliferative disorders, coagulation disorders, and nervous system
XX disorders -
XX
XX Example 1; Page 114; 119pp; English.
XX
XX The present sequence is that of Hyseq clone identification number
XX 16372272, which was obtained from a human thymus cDNA library
XX using standard PCR with primers specific for vector sequence
XX flanking the inserts, sequencing by hybridisation sequence
XX signature analysis, and Sanger sequencing techniques. This
XX expressed sequence tag was used in the assembly of a full-length
XX cDNA sequence (see AAH26343) encoding a novel human growth
XX regulatory-like polypeptide (GRLP, see AAB02671). The GRLP
XX belongs to the same protein family as growth regulatory proteins,
XX growth factors, human melanoma derived growth regulatory protein
XX precursor (64% similarity and 45% identity over 111 amino acids)
XX or melanoma inhibitory activity, cattle cartilage-derived
XX retinoic acid sensitive protein (CD-RAP, 44% identity and 64%
XX similarity over 126 amino acids) and other retinoic acid-sensitive
XX proteins. GRLP polypeptides and polynucleotides of the invention
XX can be used in the prophylaxis, treatment (including gene therapy)
XX and diagnosis of disorders and diseases caused by, or involving,
XX cartilage development and maintenance, inhibition of melanoma cell
XX growth and tumours, including neuroectodermal tumours such as
XX gliomas. The polynucleotides can also be used to design probes
XX and primers, for chromosome and gene mapping, in the recombinant
XX production of protein, in the generation of antisense, ribozyme and
XX peptide-nucleic acid molecules, and to produce transgenic animals.
XX
XX Sequence 426 BP; 119 A; 73 C; 113 G; 120 T; 1 other;
```

Query Match 100.0%; Score 330; DB 22; Length 426;
Best Local Similarity 100.0%; Pred. No. 3.2e-97;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 catgaataatttatggacgcttagcttccaagaagctctgtgcagatgatgagtgctc 60
DB 73 catgaataatttatggacgcttagcttccaagaagctctgtgcagatgatgagtgctc 132

QY 61 tatactattctctgctgctcaagaagattataatgcccgactgtagattcatt 120
DB 133 tatactattctctgctgctcaagaagattataatgcccgactgtagattcatt 192

QY 121 aacgttaaaaaaggcagcagatctatgtactcaagctgtgtaaaagaaatgagct 180
DB 193 aacgttaaaaaaggcagcagatctatgtactcaagctgtgtaaaagaaatgagct 252

QY 181 ggagaattttggcctggcagtgcttattggtggtggtggtggtggtggtggt 240
DB 253 ggagaattttggcctggcagtgcttattggtggtggtggtggtggtggtggt 312

QY 241 tatttccccaggaacttggctcaaggaacagcgtgtgtaccaggaagctaccaggaagtt 300
DB 313 tatttccccaggaacttggctcaaggaacagcgtgtgtaccaggaagctaccaggaagtt 372

QY 301 cccacacggatattgacttcttctgcgag 330
DB 373 cccacacggatattgacttcttctgcgag 402

RESULT 5
AAH98228 ID AAH98228 standard; cDNA; 891 BP.
XX AC AAH98228;
XX DT 12-OCT-2001 (first entry)
XX DE Human EST-derived coding sequence SEQ ID NO: 85.
XX KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
XX KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
XX KW diagnostics; forensic test; gene mapping; genetic disorder;
XX KW biodiversity; gene therapy; nutrition; ss.
XX OS Homo sapiens.
XX PN WO200154477-A2.
XX PD 02-AUG-2001.
XX PF 25-JAN-2001; 2001WO-US02687.
XX PR 25-JAN-2000; 2000US-0491404.
XX PR 17-JUL-2000; 2000US-0617746.
XX PR 03-AUG-2000; 2000US-0631451.
XX PR 15-SEP-2000; 2000US-0663870.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
XX PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX XX
XX DR WPI; 2001-476164/51.
XX DR P-PSDB; AAM23569.
XX PT Isolated polypeptide for treatment of diseases, diagnostics, raising
XX PT antibodies and research use -
XX PS Claim 1; Page 236; 1275pp; English.
XX CC The present invention provides the protein and coding sequences of novel
XX CC proteins from a variety of organisms, including human, dog, cat, horse,
XX CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea

urchin and tomato. These were derived from expressed sequence tags (ESTs)
from the organism of interest. They can be used in diagnostics,
forensics, gene mapping, identification of mutations, to assess
biodiversity and for nutritional purposes. The present sequence is a cDNA
of the invention.

Sequence 891 BP; 272 A; 146 C; 214 G; 259 T; 0 other;

Query Match 100.0%; Score 330; DB 22; Length 891;
Best Local Similarity 100.0%; Pred. No. 4.5e-97;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 catgaataatttatggacgcttagcttccaagaagctctgtgcagatgatgagtgctc 60
DB 73 catgaataatttatggacgcttagcttccaagaagctctgtgcagatgatgagtgctc 132

QY 61 tatactattctctgctgctcaagaagattataatgcccgactgtagattcatt 120
DB 133 tatactattctctgctgctcaagaagattataatgcccgactgtagattcatt 192

QY 121 aacgttaaaaaaggcagcagatctatgtactcaagctgtgtaaaagaaatgagct 180
DB 193 aacgttaaaaaaggcagcagatctatgtactcaagctgtgtaaaagaaatgagct 252

QY 181 ggagaattttggcctggcagtgcttattggtggtggtggtggtggtggtggt 240
DB 253 ggagaattttggcctggcagtgcttattggtggtggtggtggtggtggtggt 312

QY 241 tatttccccaggaacttggctcaaggaacagcgtgtgtaccaggaagctaccaggaagtt 300
DB 313 tatttccccaggaacttggctcaaggaacagcgtgtgtaccaggaagctaccaggaagtt 372

QY 301 cccacacggatattgacttcttctgcgag 330
DB 373 cccacacggatattgacttcttctgcgag 402

RESULT 6
AAH26342 ID AAH26342 standard; cDNA; 891 BP.
XX AC AAH26342;
XX DT 02-OCT-2001 (first entry)
XX DE Human growth regulatory-like polypeptide partial cDNA clone.
XX KW Growth regulatory-like polypeptide; human; cartilage; melanoma;
XX KW neuroectodermal tumour; glioma; cancer; gene therapy; diagnosis;
XX KW ss.
XX OS Homo sapiens.
XX PN WO200155332-A2.
XX PD 02-AUG-2001.
XX PF 25-JAN-2001; 2001WO-US02455.
XX PR 25-JAN-2000; 2000US-0491404.
XX PR 02-MAY-2000; 2000US-0563786.
XX PA (HYSE-) HYSEQ INC.
XX PI Mize NK, Boyle BJ, Ford JE, Arterburn MC, Tang YT, Liu C;
XX PI Drmanac RT;
XX DR WPI; 2001-483233/52.
XX PT Isolated human growth regulatory-like polypeptide useful for treating
XX PT e.g. Alzheimer's disease, cancer, autoimmune disorders, and
XX PT hyperproliferative disorders, coagulation disorders, and nervous system

```
PT disorders -
XX Claim 1; Page 115; 119pp; English.
XX
XX The present sequence is that of a novel nucleic acid that was
CC assembled from human thymus cDNA library-derived Hyseq clone
CC identification number 16372272 (see AAH26341). A recursive
CC algorithm was used to extend the clone by pulling additional
CC sequences from different databases. A full-length sequence (see
CC AAH26343) encoding novel human growth regulatory-like polypeptide
CC (GRLP, see AAF59083) was subsequently obtained. Human GRLP
CC belongs to the same protein family as growth regulatory proteins,
CC growth factors, human melanoma derived growth regulatory protein,
CC precursor (64% similarity and 45% identity over 111 amino acids)
CC or melanoma inhibitory activity, cattle cartilage-derived
CC retinoic acid sensitive protein (CD-RAP, 44% identity and 64%
CC similarity over 126 amino acids) and other retinoic acid-sensitive
CC proteins. GRLP polypeptides and polynucleotides of the invention
CC can be used in the prophylaxis, treatment (including gene therapy)
CC and diagnosis of disorders and diseases caused by, or involving,
CC cartilage development and maintenance, inhibition of melanoma cell
CC growth and tumours, including neuroectodermal tumours such as
CC gliomas. The polynucleotides can also be used to design probes
CC and primers, for chromosome and gene mapping, in the recombinant
CC production of protein, in the generation of antisense, ribozyme and
CC peptide-nucleic acid molecules, and to produce transgenic animals.
XX
SQ Sequence 891 BP; 272 A; 146 C; 214 G; 259 T; 0 other;

Query Match 100.0%; Score 330; DB 22; Length 891;
Best Local Similarity 100.0%; Pred. No. 4.5e-97;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 catggaattatttgaccgtctagcttccaagaagctctgtgcagatgatgagtgtgc 60
DB 73 catggaattatttgaccgtctagcttccaagaagctctgtgcagatgatgagtgtgc 132
QY 61 tatactatttctgtgctgctcaagaagattataatgccccgactgtagattcatt 120
DB 133 tatactatttctgtgctgctcaagaagattataatgccccgactgtagattcatt 192
QY 121 aacgtttaaaaaaggcagcagatctatgtgtactcaagctgtgtaaaagaaatggagct 180
DB 193 aacgtttaaaaaaggcagcagatctatgtgtactcaagctgtgtaaaagaaatggagct 252
QY 181 ggagaattttggcgtggcaggttttatggtgatggccagcagagatggagtcgtgggt 240
DB 253 ggagaattttggcgtggcaggttttatggtgatggccagcagagatggagtcgtgggt 312
QY 241 tatttccccaggaaacttggtcaaggaaacagcgtgtgtaccaggaagctaccagggaagt 300
DB 313 tatttccccaggaaacttggtcaaggaaacagcgtgtgtaccaggaagctaccagggaagt 372
QY 301 cccaccacggatattgacttcttctgcgag 330
DB 373 cccaccacggatattgacttcttctgcgag 402

RESULT 7
AAF59083
ID AAF59083 standard; DNA; 923 BP.
XX
AC AAF59083;
XX
XX
DT 23-APR-2001 (first entry)
XX
DE Human MLP; nucleotide sequence SEQ ID NO:29.
XX
XX MLP; MIA; melanoma inhibitory activity; cancer; bone disease;
KW joint disease; pathologic angiogenesis; diagnosis; antinflammatory;
KW cardiant; gene therapy; secretory cell function regulator; promoter;
KW inhibitor; ds.

XX Homo sapiens.
XX WO200102564-A1.
XX 11-JAN-2001.
XX 29-JUN-2000; 2000WO-JP04278.
XX 30-JUN-1999; 99JP-0186718.
XX (TAKE ) TAKEDA CHEM IND LTD.
XX Itoh Y, Nishi K, Ogi K, Ohkubo S, Mogi S, Noguchi Y, Yoshimura K;
XX Tanaka H;
XX WPI; 2001-159271/16.
XX Safe, low-toxicity secretory cell function-regulatory protein and
XX encoded DNA, applicable as drugs, in diagnosis and development of
XX promoters and inhibitors for preventing or treating e.g. bone and joint
XX diseases -
XX Example 1; Page 99-100; 111pp; Japanese.
XX The present invention describes novel MLP proteins and their encoding
XX DNAs. The MLP proteins and DNAs have antinflammatory and cardiant
XX activities, and can be used in gene therapy and as secretory cell
XX function regulators. The MLP proteins and DNAs can be used in drugs, in
XX the diagnosis and development of promoters and inhibitors for preventing
XX or treating bone and joint diseases as well as pathologic angiogenesis.
XX AAF59063 to AAF59099 and AAF59122 to AAF59132 represent sequences used
XX in the exemplification of the present invention.
XX
SQ Sequence 923 BP; 303 A; 147 C; 213 G; 260 T; 0 other;

Query Match 100.0%; Score 330; DB 22; Length 923;
Best Local Similarity 100.0%; Pred. No. 4.6e-97;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 catggaattatttgaccgtctagcttccaagaagctctgtgcagatgatgagtgtgc 60
DB 88 catggaattatttgaccgtctagcttccaagaagctctgtgcagatgatgagtgtgc 147
QY 61 tatactatttctgtgctgctcaagaagattataatgccccgactgtagattcatt 120
DB 148 tatactatttctgtgctgctcaagaagattataatgccccgactgtagattcatt 207
QY 121 aacgtttaaaaaaggcagcagatctatgtgtactcaagctgtgtaaaagaaatggagct 180
DB 208 aacgtttaaaaaaggcagcagatctatgtgtactcaagctgtgtaaaagaaatggagct 267
QY 181 ggagaattttggcgtggcaggttttatggtgatggccagcagagatggagtcgtgggt 240
DB 268 ggagaattttggcgtggcaggttttatggtgatggccagcagagatggagtcgtgggt 327
QY 241 tatttccccaggaaacttggtcaaggaaacagcgtgtgtaccaggaagctaccagggaagt 300
DB 328 tatttccccaggaaacttggtcaaggaaacagcgtgtgtaccaggaagctaccagggaagt 387
QY 301 cccaccacggatattgacttcttctgcgag 330
DB 388 cccaccacggatattgacttcttctgcgag 417

RESULT 8
AAH26343
ID AAH26343 standard; cDNA; 1201 BP.
XX
AC AAH26343;
XX
XX 02-OCT-2001 (first entry)
DT
```

XX DE Human growth regulatory-like polypeptide cDNA.

XX KW Growth regulatory-like polypeptide; human; cartilage; melanoma;

XX KW Neuroectodermal tumour; glioma; cancer; gene therapy; diagnosis;

XX KW ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 33..419

XX FT /tag= a

XX FT sig_peptide 33..101

XX FT /tag= b

XX FT mat_peptide 102..416

XX FT /tag= c

XX WO20015332-A2.

XX PN 02-AUG-2001.

XX PD 25-JAN-2001; 2001WO-US02455.

XX PF 25-JAN-2000; 2000US-0491404.

XX PR 02-MAY-2000; 2000US-0563786.

XX PA (HYSE-) HYSEQ INC.

XX PI Mize NK, Boyle BJ, Ford JE, Arterburn MC, Tang YT, Liu C;

XX PI Drmanac RT;

XX PI WPI; 2001-483233/52.

XX DR P-PSDB; AAB82671.

XX XX

XX PT Isolated human growth regulatory-like polypeptide useful for treating

XX PT e.g. Alzheimer's disease, cancer, autoimmune disorders, and nervous system

XX PT hyperproliferative disorders, coagulation disorders, and nervous system

XX PT disorders -

XX PS Claim 1; Page 115-116; 119pp; English.

XX CC The present sequence is that of a novel nucleic acid encoding

XX CC human growth regulatory-like polypeptide (GRLP, see AAB82671).

XX CC The sequence was assembled using human thymus cDNA library-derived

XX CC Hyseq clone identification number 16372272 (see AAB26341) as seed,

XX CC using software programs to pull additional sequences from Hyseq's

XX CC proprietary database containing expressed sequence tag sequences,

XX CC and by gel sequencing using primers to extend both 5' and 3' ends.

XX CC The predicted protein has a mol.wt. of 14 kDa unglycosylated. GRLP

XX CC belongs to the same protein family as growth regulatory proteins,

XX CC growth factors, human melanoma derived growth regulatory protein

XX CC precursor (64% similarity and 45% identity over 111 amino acids)

XX CC or melanoma inhibitory activity, cattle cartilage-derived

XX CC retinoic acid sensitive protein (CD-RAP, 44% identity and 64%

XX CC similarity over 126 amino acids) and other retinoic acid-sensitive

XX CC proteins. GRLP polypeptides and polynucleotides of the invention

XX CC can be used in the prophylaxis, treatment (including gene therapy)

XX CC and diagnosis of disorders and diseases caused by, or involving,

XX CC cartilage development and maintenance, inhibition of melanoma cell

XX CC growth and tumours, including neuroectodermal tumours such as

XX CC gliomas. The polynucleotides can also be used to design probes

XX CC and primers, for chromosome and gene mapping, in the recombinant

XX CC production of protein, in the generation of antisense, ribozyme and

XX CC peptide-nucleic acid molecules, and to produce transgenic animals.

XX CC They may also have cytokine and cell proliferation or

XX CC differentiation activity, stem cell growth factor activity,

XX CC haematopoiesis regulating activity, tissue growth activity,

XX CC immunosuppressive or immunostimulant activity, activin/inhibin

XX CC activity, chemotactic/chemokinetic activity, haemostatic and

XX CC thrombolytic activity, use in cancer diagnosis and therapy,

XX CC drug screening, receptor/ligand activity, antiinflammatory

XX CC activity, and treatment of leukaemia, nervous system disorders,

XX CC arthritis and inflammation.

XX SQ Sequence 1201 BP; 357 A; 188 C; 275 G; 381 T; 0 other;

Query Match 100.0%; Score 330; DB 22; Length 1201;

Best Local Similarity 100.0%; Pred. No. 5.le-97;

Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 catgaataattatggaccgtctagctccaagaagctctgtgcagatgaatgagtgctc 60

DB 87 catgaataattatggaccgtctagctccaagaagctctgtgcagatgaatgagtgctc 146

QY 61 tatactatttctctggctagctcaagaagattataatgccccgactgtagattcatt 120

DB 147 tatactatttctctggctagctcaagaagattataatgccccgactgtagattcatt 206

QY 121 aacgttaaaaaaggcgagcagatctatgttactcaaaagctggtataaagaatgagct 180

DB 207 aacgttaaaaaaggcgagcagatctatgttactcaaaagctggtataaagaatgagct 266

QY 181 ggagaattttggctggcagtggtttatgtggtggtggtggtggtggtggtggtggtggt 240

DB 267 ggagaattttggctggcagtggtttatgtggtggtggtggtggtggtggtggtggt 326

QY 241 tatttccccaggaacttggtcaaggaacagcgtgtgtaccaggaagctaccaggaagtt 300

DB 327 tatttccccaggaacttggtcaaggaacagcgtgtgtaccaggaagctaccaggaagtt 386

QY 301 cccaccacggtatgacttcttctgcgag 330

DB 387 cccaccacggtatgacttcttctgcgag 416

RESULT 9

AAF59080

ID AAF59080 standard; DNA; 330 BP.

XX AC AAF59080;

XX DT 23-APR-2001 (first entry)

XX DE Mouse MLP nucleotide sequence SEQ ID NO:25.

XX KW MLP; MTA; melanoma inhibitory activity; cancer; bone disease;

XX KW joint disease; pathologic angiogenesis; diagnosis; antiinflammatory;

XX KW cardiant; gene therapy; secretory cell function regulator; promoter;

XX KW inhibitor; ds.

XX OS Mus musculus.

XX PN WO200102564-A1.

XX PD 11-JAN-2001.

XX PF 29-JUN-2000; 2000WO-JP04278.

XX PR 30-JUN-1999; 99JP-0186718.

XX PA (TAKE) TAKEDA CHEM IND LTD.

XX PI Itoh Y, Nishi K, Ogi K, Ohkubo S, Mogi S, Noguchi Y, Yoshimura K;

XX PI Tanaka H;

XX DR WPI; 2001-159271/16.

XX DR P-PSDB; AAB69127.

XX PT Safe, low-toxicity secretory cell function-regulatory protein and

XX PT encoded DNA, applicable as drugs, in diagnosis and development of

XX PT promoters and inhibitors for preventing or treating e.g. bone and joint

XX PT diseases -

XX PS Claim 10; Page 98; 111pp; Japanese.

XX XX


```
XX (TAKE ) TAKEDA CHEM IND LTD.
PA
XX Itoh Y, Nishi K, Ogi K, Ohkubo S, Mogi S, Noguchi Y, Yoshimura K;
PI Tanaka H;
XX WPI: 2001-159271/16.
XX
XX Safe, low-toxicity secretory cell function-regulatory protein and
PT encoded DNA, applicable as drugs, in diagnosis and development of
PT promoters and inhibitors for preventing or treating e.g. bone and joint
PT diseases -
XX
XX Example 2; Page 100-101; 111pp; Japanese.
PS
XX The present invention describes novel MLP proteins and their encoding
XX DNAs. The MLP proteins and DNAs have antiinflammatory and cardiant
CC activities, and can be used in gene therapy and as secretory cell
CC function regulators. The MLP proteins and DNAs can be used in drugs, in
CC the diagnosis and development of promoters and inhibitors for preventing
CC or treating bone and joint diseases as well as pathologic angiogenesis.
CC AAF59063 to AAF59099 and AAB69122 to AAB69132 represent sequences used
CC in the exemplification of the present invention.
XX
XX Sequence 947 BP; 279 A; 158 C; 221 G; 289 T; 0 other;
SQ
Query Match 76.4%; Score 252.2; DB 22; Length 947;
Best Local Similarity 85.4%; Pred. No. 9.2e-72;
Matches 281; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
QY 1 catggaattttatggaccctctagcttccaagaagctctgtgcagatgatgagtgtgc 60
DB 65 catggtgtattatggataaaactttcttctaagaagtgtgtgcgagtgagggtgtgc 124
QY 61 tatactatttcttgctgctcgaagagattataatgccccgactgtgattcatt 120
DB 125 tatactatttcttgccaagagcacaggaagtatactatgccagactgtaggttcac 184
QY 121 aacgttaaaaaaggcagcagatctatgttactcaaaagctggttaaaagaatggagct 180
DB 185 gagtcaagaagaaggcagcagatctatgttactccaagctggttaacagaacaggagct 244
QY 181 ggagaattttggctggcagtggtttatgtgatggccaggaagagatgggagtcgtgggt 240
DB 245 ggagagttttggctggcagtggtttatgttgaccaccaggatgagatgggaaattgaggt 304
QY 241 tatttcccaggaactgttcaaggaacacagctgtgtaccaggaagctaccaggaagtt 300
b 305 tatttcccagcaactgttggaaggagcagctgtataccadgaggcccaaggagatc 364
QY 301 cccaccacggatattgacttcttcttcgca 329
DB 365 ccaaccacggatattgacttcttcttcgta 393
RESULT 12
AAF59099
ID AAF59099 standard; DNA; 330 BP.
XX
XX AAF59099;
AC
XX
XX 23-APR-2001 (first entry)
DT
XX Rat MLP nucleotide sequence SEQ ID NO:48.
DE
XX MLP; MTA; melanoma inhibitory activity; cancer; bone disease.
KW joint disease; pathologic angiogenesis; diagnosis; antiinflammatory;
KW cardiant; gene therapy; secretory cell function regulator; promoter;
KW inhibitor; ds.
XX
XX Rattus sp.,
OS
XX
```

```
PN W0200102564-A1.
XX
XX 11-JAN-2001.
XX
XX 29-JUN-2000; 2000WO-JP04278.
XX
XX 30-JUN-1999; 99JP-0186718.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Itoh Y, Nishi K, Ogi K, Ohkubo S, Mogi S, Noguchi Y, Yoshimura K;
PI Tanaka H;
XX
XX WPI: 2001-159271/16.
DR
XX P-PSDB; AAB69131.
DR
XX Safe, low-toxicity secretory cell function-regulatory protein and
PT encoded DNA, applicable as drugs, in diagnosis and development of
PT promoters and inhibitors for preventing or treating e.g. bone and joint
PT diseases -
XX
XX Claim 12; Page 107; 111pp; Japanese.
PS
XX The present invention describes novel MLP proteins and their encoding
XX DNAs. The MLP proteins and DNAs have antiinflammatory and cardiant
CC activities, and can be used in gene therapy and as secretory cell
CC function regulators. The MLP proteins and DNAs can be used in drugs, in
CC the diagnosis and development of promoters and inhibitors for preventing
CC or treating bone and joint diseases as well as pathologic angiogenesis.
CC AAF59063 to AAF59099 and AAB69122 to AAB69132 represent sequences used
CC in the exemplification of the present invention.
XX
XX Sequence 330 BP; 91 A; 62 C; 91 G; 86 T; 0 other;
SQ
Query Match 75.5%; Score 249; DB 22; Length 330;
Best Local Similarity 84.8%; Pred. No. 6.2e-71;
Matches 279; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 1 catggaattttatggaccctctagcttccaagaagctctgtgcagatgatgagtgtgc 60
DB 1 catggtcatgtttatggataaaactttcttctaagaagtgtgtgcagatgaggagtggtgc 60
QY 61 tatactatttcttgctgctcgaagattataatgccccgactgtgattcatt 120
DB 61 tatactatttcttgccaagagcacaggaagtatactcaatgcccgactgtgattcatt 120
QY 121 aacgttaaaaaaggcagcagatctatgttactcaaaagctggttaaaagaatggagct 180
DB 121 aatgtcaagaagaaggcagcagatctatgtttatttccaagctggttaacagaacaggagct 180
QY 181 ggagaattttggctggcagtggtttatgtgatggccaggaagagatgggagtcgtgggt 240
DB 181 ggaggcaattcttgctggcagtggtttatgttgaccaccaggatgagatgggaaattgaggt 240
QY 241 tatttcccaggaactgttcaaggaacacagctgtgtaccaggaagctaccaggaagtt 300
DB 241 tatttcccagcaactgtttagagagcaacgagtgtagcaggaggccaccaaggagatt 300
QY 301 cccaccacggatattgacttcttcttcgca 329
DB 301 ccaaccacggatattgacttcttcttcgta 329
RESULT 13
AAF59098
ID AAF59098 standard; DNA; 384 BP.
XX
XX AAF59098;
AC
XX
XX 23-APR-2001 (first entry)
DT
XX Rat MLP nucleotide sequence SEQ ID NO:46.
DE
```



```
XX MLP; MIA; melanoma inhibitory activity; cancer; bone disease;
KW joint disease; pathologic angiogenesis; diagnosis; antinflammatory;
KW cardiant; gene therapy; secretory cell function regulator; promoter;
KW inhibitor; ds.
XX
OS Rattus sp.
XX
XX WO200102564-A1.
XX
XX 11-JAN-2001.
XX
XX 29-JUN-2000; 2000WO-JP04278.
XX
XX 30-JUN-1999; 99JP-0186718.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Itoh Y, Nishi K, Ogi K, Ohkubo S, Mogi S, Noguchi Y, Yoshimura K;
XX Tanaka H;
XX
XX WPI; 2001-159271/16.
XX
XX P-PSDB; AAB69130.
XX
XX Safe, low-toxicity secretory cell function-regulatory protein and
XX encoded DNA, applicable as drugs, in diagnosis and development of
XX promoters and inhibitors for preventing or treating e.g. bone and joint
XX diseases
XX
XX Claim 13; Page 105-106; 111pp; Japanese.
XX
XX The present invention describes novel MLP proteins and their encoding
XX DNAs. The MLP proteins and DNAs have antinflammatory and cardiant
XX activities, and can be used in gene therapy and as secretory cell
XX function regulators. The MLP proteins and DNAs can be used in drugs, in
XX the diagnosis and development of promoters and inhibitors for preventing
XX or treating bone and joint diseases as well as pathologic angiogenesis.
XX AAF59063 to AAF59099 and AAB69122 to AAB69132 represent sequences used
XX in the exemplification of the present invention.
XX
XX Sequence 384 BP; 98 A; 72 C; 109 G; 105 T; 0 other;

Query Match 75.5%; Score 249; DB 22; Length 384;
Best Local Similarity 84.8%; Pred. No. 6.7e-71;
Matches 279; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

1 catgggaattattatggaccgtctagcttccaagaagctctgtgcagatgatgagtggtgc 60
||||| |||||||| || ||||| ||||||| ||||||| ||||||| |||||||
55 catggcatgtttatgataaactttcttctaagaagtgtgtgcagatgagagtggtgc 114

61 tatactattctctggtcagtgctcaagaagattataatgcccgactgtagattcatt 120
||||| |||||||| || ||||| ||||||| ||||||| ||||||| |||||||
115 tatactattctctggtcagtgctcaagaagattataatgcccgactgtagattcatt 174

121 aacgtttaaaaaaggccagcagatctatgttactcaaaagctgtaaaaaaaggagct 180
||||| |||||||| ||||||| ||||||| ||||||| ||||||| |||||||
175 aatgtcaagaagaaggccagcagatctatgttactcaaaagctgtaaaaaaaggagct 234

181 ggagaattttggcgtggcagtgcttattgtgtgagcagacagatggagtcgtgggt 240
||||| |||||||| ||||||| ||||||| ||||||| ||||||| |||||||
235 ggagcattctggcgtggcagtgcttattgtgtgagcagacagatggagtcgtgggt 294

241 tatttccccaggaacttggtcaaggacacagcgtgtgtaccaggaagctaccaggaggt 300
||||| |||||||| ||||||| ||||||| ||||||| ||||||| |||||||
295 tatttccccaggaacttggttagagagcaacagcagtgtagcaggaggccaccaggagatt 354

301 ccaacacagatattgacttcttctcga 329
|| ||||||||| ||||||||| ||
355 ccaacacagatattgacttcttcttga 383
|| ||||||||| ||||||||| ||
```

RESULT 14

```
AAF59093
ID AAF59093 standard; DNA; 307 BP.
XX
XX AAF59093;
XX
XX 23-APR-2001 (first entry)
XX
XX Rat MLP nucleotide sequence SEQ ID NO:41.
XX
XX MLP; MIA; melanoma inhibitory activity; cancer; bone disease;
XX joint disease; pathologic angiogenesis; diagnosis; antinflammatory;
XX cardiant; gene therapy; secretory cell function regulator; promoter;
XX inhibitor; ds.
XX
XX Rattus sp.
XX
XX WO200102564-A1.
XX
XX 11-JAN-2001.
XX
XX 29-JUN-2000; 2000WO-JP04278.
XX
XX 30-JUN-1999; 99JP-0186718.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Itoh Y, Nishi K, Ogi K, Ohkubo S, Mogi S, Noguchi Y, Yoshimura K;
XX Tanaka H;
XX
XX WPI; 2001-159271/16.
XX
XX Safe, low-toxicity secretory cell function-regulatory protein and
XX encoded DNA, applicable as drugs, in diagnosis and development of
XX promoters and inhibitors for preventing or treating e.g. bone and joint
XX diseases
XX
XX Example 9; Page 104; 111pp; Japanese.
XX
XX The present invention describes novel MLP proteins and their encoding
XX DNAs. The MLP proteins and DNAs have antinflammatory and cardiant
XX activities, and can be used in gene therapy and as secretory cell
XX function regulators. The MLP proteins and DNAs can be used in drugs, in
XX the diagnosis and development of promoters and inhibitors for preventing
XX or treating bone and joint diseases as well as pathologic angiogenesis.
XX AAF59063 to AAF59099 and AAB69122 to AAB69132 represent sequences used
XX in the exemplification of the present invention.
XX
XX Sequence 307 BP; 84 A; 56 C; 90 G; 77 T; 0 other;
```

Query Match 65.0%; Score 214.6; DB 22; Length 307;
Best Local Similarity 82.1%; Pred. No. 9.3e-60;
Matches 247; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

```
QY 1 catgggaattattatggaccgtctagcttccaagaagctctgtgcagatgatgagtggtgc 60
||||| |||||||| || ||||| ||||||| ||||||| ||||||| |||||||
Db 7 catgggtattattatggataaactttcttctaagaagtgtgtgcagatgagagtggtgc 66

61 tatactattctctggtcagtgctcaagaagattataatgcccgactgtagattcatt 120
||||| |||||||| || ||||| ||||||| ||||||| ||||||| |||||||
Db 67 tatactattctctggtcagtgctcaagaagattataatgcccgactgtagattcatt 126

121 aacgtttaaaaaaggccagcagatctatgttactcaaaagctgtaaaaaaaggagct 180
||||| |||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 127 aatgtcaagaagaaggccagcagatctatgttactcaaaagctgtaaaaaaaggagct 186

181 ggagaattttggcgtggcagtgcttattgtgtgagcagacagatggagtcgtgggt 240
||||| |||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 187 ggagcattctggcgtggcagtgcttattgtgtgagcagacagatggagtcgtgggt 246

241 tatttccccaggaacttggtcaaggacacagcgtgtgtaccaggaagctaccaggaggt 300
||||| |||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 247 tatttccccaggaacttggttagagagcaacagcagtagtaccaggaggccaccaggagat 306
```

QY 301 c 301
|
Db 307 c 307

RESULT 15
AAF59092
ID AAF59092 standard; DNA; 261 BP.

XX AAF59092;

AC AAF59092;

DT 23-APR-2001 (first entry)

XX Rat MLP nucleotide sequence SEQ ID NO:40.

XX MLP; MIA; melanoma inhibitory activity; cancer; bone disease;
KW joint disease; pathologic angiogenesis; diagnosis; antiinflammatory;
KW cardiant; gene therapy; secretory cell function regulator; promoter;
KW inhibitor; ds.

XX Rattus sp.

XX WO200102564-A1.

PN 11-JAN-2001.

XX 29-JUN-2000; 2000WO-JP04278.

XX 30-JUN-1999; 99JP-0186718.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Itoh Y, Nishi K, Ogi K, Ohkubo S, Mogi S, Noguchi Y, Yoshimura K;
PI Tanaka H;

XX WPI; 2001-159271/16.

XX Safe, low-toxicity secretory cell function-regulatory protein and
PT encoded DNA, applicable as drugs, in diagnosis and development of
PT promoters and inhibitors for preventing or treating e.g. bone and joint
PT diseases

XX Example 9; Page 104; 111pp; Japanese.

XX The present invention describes novel MLP proteins and their encoding
CC DNAs. The MLP proteins and DNAs have antiinflammatory and cardiant
CC activities, and can be used in gene therapy and as secretory cell
CC function regulators. The MLP proteins and DNAs can be used in drugs, in
CC the diagnosis and development of promoters and inhibitors for preventing
CC or treating bone and joint diseases as well as pathologic angiogenesis.
CC AAF59063 to AAF59099 and AAB69122 to AAB69132 represent sequences used
CC in the exemplification of the present invention.

XX Sequence 261 BP; 72 A; 46 C; 75 G; 68 T; 0 other;

Query Match 58.7%; Score 193.8; DB 22; Length 261;
Best Local Similarity 83.9%; Pred. No. 5e-53;
Matches 219; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 15 ggaccgtctagcttccaaagaagctctgtgcagatgatgagtggtctatactattctct 74
Db 1 ggataaacttcttccaaagaagtggtgcagatgagtggtctatactattctct 60

QY 75 ggctagtgtcgaagattataatgcccgcactgtagcttaccattacgttaaaaaagg 134
Db 61 ggcaagagcaacaggaagactacaatgcccggactgtggttccatcatgtcaagaaagg 120

QY 135 gcagcagatctatgtgtactaaagctgttaaaagaaatggagctggagaattttgggc 194
Db 121 gcagcagatctatgtttattccaagctgttaacagaaaaatggagctgggcattctctgggc 180

QY 195 tggcagtggtttatggtgatggccagcagacgagatgggagtcgtgggttatttccccagga 254
|
Db 181 tggcagtggtttatggtgatggccagcagacgagatgggagtcgtgggttatttccccagga 240
QY 255 cttggtcaagggaacacgcgtgt 275
|
Db 241 cttggttagagagcaacgagt 261

Search completed: September 23, 2002, 15:55:37
Job time: 1146 sec

THIS PAGE BLANK (USPTO)

EMBO J. 8, 2283-2290, 1989

[illegible]

C:Accession: T32734
R:Lee, W.L.; Ostap, E.M.; Zot, H.G.; Pollard, T.D.
submitted to the EMBL Data Library, August 1998
A:Description: Hydrodynamic and ligand binding properties of Acanthamoeba myosin-IA G
A:Reference number: Z21216
A:Accession: T32734
A:Status: preliminary; translated from GB/EMBL/DDBB
A:Molecule type: DNA
A:Residues: 1-1215 <LEE>
A:Cross-references: EMBL:AF085185; NID:g3599477; PIDN:AAC35357.1
A:Experimental source: strain Neff
C:Genetics:
A:Gene: MIA
C:Superfamily: protozoan myosin heavy chain IB; myosin motor domain homology; SH3 hom
F:14-674/Domain: myosin motor domain homology <MWO>

Query Match 12.68; Score 74.5; DB 2; Length 1215;
Best Local Similarity 28.6%; Pred. No. 18;
Matches 20; Conservative 12; Mismatches 25; Indels 13; Gaps 2;

QY 20 VYTISLASAQDYNAPDCRFNVAKGOIYVYSKLVKENGAGFEWAGSVYGDQDEMGVV 79
| | | | | : : : | : : | | : | | : |
Db 1158 VPTVGRCALYDYGAEADETLREGDVIDVQK-----SGEWWESTLNGK-----T 1204

QY 80 GYFPRNLVKE 89
| | | | | : : : | : : | | : | | : |
Db 1205 GVFPANVYED 1214

RESULT 7
T48525
hypothetical protein T22P22.50 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T48525
R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancro
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24490
A:Accession: T48525
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-308 <BEV>
A:Cross-references: EMBL:AL163814
A:Experimental source: cultivar Columbia; BAC clone T22P22
C:Genetics:
A:Map position: 5
A:Introns: 63/3; 117/3; 135/3; 180/3
A:Note: T22P22.50

Query Match 12.5%; Score 74; DB 2; Length 308;
Best Local Similarity 37.2%; Pred. No. 4.5;
Matches 16; Conservative 8; Mismatches 17; Indels 2; Gaps 1;

QY 44 KGQYTVYSKLVKENGAGFEWAGSVYGDQDEMGVVGYFPNRL 86
| | | | | : | : | | | | | : | | : |
Db 168 KGQERAVYASVQEKDIGNWVTDVRYDRPDVH--ICGWPKEL 208

RESULT 8
C64439
asparagine synthase (glutamine-hydrolyzing) (EC 6.3.5.4) MJ1116 [similarity] - Methan
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
C:Accession: C64439
R:Bult, C.J.; White
C:Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

Db 35 YLQLLAQKYDCEEKVVTETIINLKA-----ILNLPKGTETHEFVSDLHGEYQAFQHV 83

Db 35 YLQLLAQKYDCEEKVVTETIINLKA-----ILNLPKGT EHFVSDLHG EYQAFQHV 83

RESULT 12
C90186
AAA family ATPase [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C:Accession: C90186
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: C90186
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-769 <KUR>
A:Cross-References: GB:AE006641; NID:gl13813572; PIDN:AAK40746.1; GSPDB:GN00155
C:Genetics:
A:Gene: SSO0421
Superfamily: transitional endoplasmic reticulum ATPase; FtsH/SEC18/CDC48-type ATP-bind

Query Match 12.0%; Score 71; DB 2; Length 769;
Best Local Similarity 35.0%; Pred. No. 25;
Matches 28; Conservative 10; Mismatches 20; Indels 22; Gaps 5;
QY 49 YVSKLVKENG-----AGEFNAGSVYGDG--QDEMVGVGFPNRL-----VKE 89
DB 51 YTISRLGIENGDIYVEIIGPSGSAQAALIGDGIADNEIRVDGYIRSGVIGIGDEVTVKR 110
QY 90 ORVYQEAQKEV--PTTDIDF 107
DB 111 AQV-QDATKVLAPTOPISF 129

RESULT 13
D72278
endo-1,4-beta-mannosidase - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: D72278
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316
C:Accession: D72278
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-669 <ARN>
A:Cross-References: GB:AE001779; GB:AE000512; NID:g4981777; PIDN:AAD36302.1; PID:g498178
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1227

Query Match 11.9%; Score 70.5; DB 2; Length 669;
Best Local Similarity 27.6%; Pred. No. 24;
Matches 32; Conservative 12; Mismatches 37; Indels 35; Gaps 6;
QY 25 LASAQEDYNAPDCRFINVKKGQOI-----YVYSK-----LVKENGAGEF 63
DB 393 LFTGTGDIREDTCSFLPLPKDGMIEKKTVEVRAGVFDYSNTEKLSVKYKEDLVFENEIHL 452
QY 64 WAGSVYV-----DGDQEMGVGYFPRNLVKEQ---RVYQEAQKEVPTTDIDF 107
DB 453 GYG-IGFPLDTRTPDGEHEMFLEGHFGKIVKSIKAKVVNEA-RVVLAEVDV 506

RESULT 14
D86758
orotate phosphoribosyltransferase (EC 2.4.2.10) [imported] - Lactococcus lactis subsp. J

C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: D86758
R:Botolin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Eh
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: D86758
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-209 <STO>
A:Cross-References: GB:AE005176; PID:gl12724022; PIDN:AAK05166.1; GSPDB:GN00146
A:Experimental source: strain ILL403
C:Genetics:
A:Gene: pyrE
C:Keywords: glycosyltransferase; pentosyltransferase

Query Match 11.9%; Score 70; DB 2; Length 209;
Best Local Similarity 26.9%; Pred. No. 7.7;
Matches 21; Conservative 15; Mismatches 42; Indels 0; Gaps 0;
QY 25 LASAQEDYNAPDCRFINVKKGQOIYVYSLVKENGAGEFAGSVYGDQDEMVGVGFPFR 84
DB 95 IRSKPKDHGAGQVEGRVTGQKMWVVEDLISTGGSVLEAVAAAREGADVLGVVAIFTY 154
QY 85 NLVKQQRVYQEAQKEVPT 102
DB 155 ELEKANRKEPADAGVKLAT 172

RESULT 15
F90113
cell division cycle 2 homolog [imported] - Guillardia theta nucleomorph
C:Species: nucleomorph Guillardia theta
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 15-Jun-2001
C:Accession: F90113
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.;
Nature 410, 1091-1096, 2001
A:Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A99082; MUID:11323671
A:Accession: F90113
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-295 <DOU>
A:Cross-References: GB:AJ010592; NID:gl2580757; PIDN:CAC27075.1; GSPDB:GN00151
C:Genetics:
A:Map position: 2
A:Genome: nucleomorph
C:Superfamily: Kinase-related transforming protein; protein kinase homology
C:Keywords: nucleomorph

Query Match 11.8%; Score 69.5; DB 2; Length 295;
Best Local Similarity 28.0%; Pred. No. 13;
Matches 28; Conservative 14; Mismatches 31; Indels 27; Gaps 6;
QY 8 LASKKLCADECYVITISLASAQEDYNAPDCRFINVKKGQOIYVYSLVKENGAGEF 65
DB 157 LCSKKICLSKIV-TLW-----YRAPE-----ILLGQHFYDYSVDMWSFGCVIGELIT 203
QY 66 GSVYGDQDEMVGVGFPNRLVKEQQRVYQEAQKEVPTTDI 105
DB 204 GEILFOGKSELQDL-----NKIFO--TIGTPTTEI 231

Search completed: September 23, 2002, 09:37:17
Job time: 27 sec

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	226.5	38.4	96	11	Q9J109	Q9j109 mesocricetu	
2	185.5	31.4	119	4	Q96PC5	Q96pc5 homo sapien	
3	181	30.7	137	11	Q921X3	Q921x3 mus musculu	
4	171.5	29.1	119	11	Q91ZV0	Q91zv0 mus musculu	
5	168.5	28.6	88	13	Q90XF1	Q90xf1 tetraodon n	
6	88.5	15.0	1196	4	Q9OH2	Q9oh2 homo sapien	
7	86.5	14.7	719	4	Q96D37	Q96d37 homo sapien	
8	85.5	14.5	166	11	Q08526	Q08526 mus musculu	
9	82.5	14.0	498	4	Q9H803	Q9h803 homo sapien	
10	77	13.1	331	2	Q93J77	Q93j77 pseudomonas	
11	74.5	12.6	1215	5	Q97202	Q97202 acanthamoeb	
12	74	12.5	308	10	Q9LXG4	Q9lyg4 arabidopsis	
13	74	12.5	2161	4	Q9Y566	Q9y566 homo sapien	
14	73	12.4	636	10	Q9FTZ8	Q9ftz8 oryza sativ	
15	73	12.4	643	10	Q9FU04	Q9fu04 oryza sativ	
16	72	12.2	344	10	Q9FU07	Q9fu07 oryza sativ	


```
Q9H0H2
ID Q9H0H2 PRELIMINARY; PRT; 1196 AA.
AC Q9H0H2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 137.1 KDA PROTEIN.
GN DKF2P3434N031.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansong W., Boecker M., Blocker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA Meves H.W., Ottenvaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a Catalog of Human Genes and Proteins: Sequencing and
RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";
RL Genome Res. 11:422-435(2001).
- !- SIMILARITY: CONTAINS 4 WD REPEATS (TRP-ASP DOMAINS).
CC EMBL; AL136797; CAB66731.1;
DR HSSP; P06241; 1SWF
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00400; WD40; 4.
DR PRINTS; PR00452; SH3DOMAIN.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00320; WD40; 6.
DR PROSITE; PS50002; SH3; 1.
DR PROSITE; PS50082; WD_REPEATS_2; 4.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Repeat; WD repeat.
SQ SEQUENCE 1196 AA; 137114 MW; E27E3FB1D7932024 CRC64;

Query Match 15.0%; Score 88.5; DB 4; Length 1196;
Best Local Similarity 33.3%; Pred. No. 0.57;
Matches 27; Conservative 12; Mismatches 31; Indels 11; Gaps 4;

QY 20 VYTIASAEQEDYNAPDCRFINVKKQQIYVYVKLVKENGAGFEWAGSVYGDQDEGVV 79
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +
1050 VDPAPTVALYDYTNARSDSLTHRGDIIRVF---FKDN--EDWYGSI-GKQGE----- 1098
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +

80 GYFPLNVLKQEVQATKEV 100
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +
1099 GYFPAHVASETLYQELPPEI 1119
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +

Query Match 14.5%; Score 85.5; DB 11; Length 166;
Best Local Similarity 31.4%; Pred. No. 0.1;
Matches 22; Conservative 14; Mismatches 21; Indels 13; Gaps 3;

QY 26 ASAEQEDYNAPDCRFINVKKQQIYVYVKLVKENGAGFEWAGSVYGDQDEGVVGYFPRN 85
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +
108 AKARYDFCARDSELSLKEGDII----KILNKKQQGQGWGRGEIYGR-----IGWFPSPN 156
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +

86 LVKEQRYQVE 95
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +
157 YVEED--YSE 164
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +

RESULT 9
Q9H803 PRELIMINARY; PRT; 498 AA.
ID Q9H803
AC Q9H803;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
KW Hypothetical protein.
```

```
SQ SEQUENCE 719 AA; 83727 MW; A51B757DA543BA6C CRC64;

Query Match 14.7%; Score 86.5; DB 4; Length 719;
Best Local Similarity 32.9%; Pred. No. 0.5;
Matches 23; Conservative 13; Mismatches 21; Indels 13; Gaps 3;

QY 26 ASAEQEDYNAPDCRFINVKKQQIYVYVKLVKENGAGFEWAGSVYGDQDEGVVGYFPRN 85
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +
661 AKARYDFCARDSELSLKEGDII----KILNKKQQGQGWGRGEIYGR-----VGWFPAN 709
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +

86 LVKEQRYQVE 95
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +
710 YVEED--YSE 717
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +

RESULT 8
Q9H803 PRELIMINARY; PRT; 166 AA.
ID Q9H803;
AC Q9H803;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VAV-T.
GN VAV.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DDY; TISSUE=TESTIS;
RX MEDLINE=97190224; PubMed=9038379;
RA Okumura K., Kaneko Y., Nonoguchi K., Nishiyama H., Yokoi H.,
RA Higuchi T., Itoh K., Yoshida O., Miki T., Fujita J.;
RT "Expression of a novel isoform of Vav, Vav-T, containing a single
RT Src homology 3 domain in murine testicular germ cells.";
RL Oncogene 14:713-720(1997).
DR EMBL; D83286; BAA18950.1; -.
DR HSSP; Q60631; IGBO.
DR MGD; MGI:98923; Vav.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
SQ SEQUENCE 166 AA; 19434 MW; 2C3AF1FF07A5018F CRC64;

Query Match 14.5%; Score 85.5; DB 11; Length 166;
Best Local Similarity 31.4%; Pred. No. 0.1;
Matches 22; Conservative 14; Mismatches 21; Indels 13; Gaps 3;

QY 26 ASAEQEDYNAPDCRFINVKKQQIYVYVKLVKENGAGFEWAGSVYGDQDEGVVGYFPRN 85
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +
108 AKARYDFCARDSELSLKEGDII----KILNKKQQGQGWGRGEIYGR-----IGWFPSPN 156
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +

86 LVKEQRYQVE 95
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +
157 YVEED--YSE 164
+ + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + + +

RESULT 9
Q9H803 PRELIMINARY; PRT; 498 AA.
ID Q9H803
AC Q9H803;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
KW Hypothetical protein.
```

```
DE HYPOTHETICAL 57.0 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYO;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugeno S., Shiratori A., Sudo H.,
RA Takatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 4 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL; AK024085; BAB14820.1; -.
DR HSSP; P06241; LSHP.
DR InterPro; IPR002453; Beta_tubulin.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00400; WD40; 4.
DR PRINTS; PR00320; SH3DOMAIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS50002; SH3; 1.
DR PROSITE; PS00228; TUBULIN_B_AUTOREG; UNKNOWN_1.
DR PROSITE; PS50082; WD_REPEATS_2; 4.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 498 AA; 57011 MW; 0D2F58A28A4553AC CRC64;

Query Match 14.0%; Score 82.5; DB 4; Length 498;
Best Local Similarity 32.1%; Pred. No. 0.89;
Matches 26; Conservative 12; Mismatches 32; Indels 11; Gaps 4;

QY 20 VYTIASLAQEDYNAPDCRFINVKKGQIYVYVSKLVKENGAGFEWAGSVYGGQDGMGVV 79
Db 404 VDTAPTVVALDYTNARSDELIRHGDIIKRVF---FKDN--EDWYVGSIGKQGE----- 452

QY 80 GYFPRNLVKRQRYQATKEV 100
Db 453 GYLPANHVASETLYQELPPEI 473

RESULT 10
Q93JT7 PRELIMINARY; PRT; 331 AA.
AC Q93JT7;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DE PROBABLE C4 DICARBOXYLATE BINDING PROTEIN.
GN DCTP.
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=316;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A15;
RA Lin M., Gao L., Desnoes N., Elmerich C.;
RT "Dicarboxylic acid transport in Pseudomonas stutzeri.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ313422; CAC44171.1; -.
SQ SEQUENCE 331 AA; 36729 MW; CB893C98508BEC39 CRC64;
```

```
Query Match 13.1%; Score 77; DB 2; Length 331;
Best Local Similarity 25.6%; Pred. No. 2.2;
Matches 30; Conservative 22; Mismatches 43; Indels 22; Gaps 5;

QY 5 MDRLASKKLCADECVYTISLASAQEDYNAPDCRFI-----NVKKGQIYVYVSKLVKENG 59
Db 1 MFKLTAkala----CALSLSIAGLAHADPITIKFSHVVAENTPKGGALMFKKLVLEERL 56

QY 60 AGE-----FWAGSVYGGDGE-----MGVYGYPRNLVKRQRYQATKEVPTTIDIF 107
Db 57 AGKVEQVYVPSNSLFGDGKREALLGLDYLTIAPSLAK-----FEHYSKGVQVFDLPF 109

RESULT 11
O77202 PRELIMINARY; PRT; 1215 AA.
AC O77202;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE MYOSIN-IA.
GN MIA.
OS Acanthamoeba castellanii (Amoeba).
OC Eukaryota; Acanthamoebidae; Acanthamoeba.
OX NCBI_TaxID=5755;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEFF;
RA Lee W.-L., Ostap E.M., Zot H.G., Pollard T.D.;
RT "Hydrodynamic and ligand binding properties of Acanthamoeba Myosin-IA
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF085185; AAC35357.1; -.
DR HSSP; P08799; 1MND.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00242; MYSC; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
SQ SEQUENCE 1215 AA; 134103 MW; FEFBFCC2EF936C72 CRC64;

Query Match 12.6%; Score 74.5; DB 5; Length 1215;
Best Local Similarity 28.6%; Pred. No. 21;
Matches 20; Conservative 12; Mismatches 25; Indels 13; Gaps 2;

QY 20 VYTIASLAQEDYNAPDCRFINVKKGQIYVYVSKLVKENGAGFEWAGSVYGGQDGMGVV 79
Db 1158 VPTVGRCALYDYGAEDELTLREGDVIDVIQK-----SGEWWEGLNGK-----T 1204

QY 80 GYFPRNLVKR 89
Db 1205 GVFPANYVED 1214

RESULT 12
Q9LYG4 PRELIMINARY; PRT; 308 AA.
AC Q9LYG4;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DE HYPOTHETICAL 34.0 KDA PROTEIN.
GN T22P22_50.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
```

[illegible]

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.

OX NCBI_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. NIPPONBARE;

RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC

RT clone:P0698G03.;"

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. NIPPONBARE;

RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC

RT clone:P0494A10.;"

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF002747; BAB17342.1; -

DR EMBL; AF002541; BAB55464.1; -

DR HSSP; P08631; IADS

DR InterPro; IPR000719; Euk_pkinase.

DR InterPro; IPR002290; Ser_thr_pkinase.

DR InterPro; IPR004040; STY_pkinase.

DR InterPro; IPR001245; Tyr_pkinase.

DR Pfam; PF00069; pkinase; 1.

DR SMART; SM00221; STYK; 1.

DR SMART; SM00220; S_TKG; 1.

DR SMART; SM00219; TYRK; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.

KW ATP-binding; Kinase; Receptor; Transferase.

SQ SEQUENCE 643 AA; 71525 MW; BFB45CB172A6F15A CRC64;

Query Match

Best Local Similarity 12.4%; Score 73; DB 10; Length 643;

Matches 31; Conservative 15; Mismatches 52; Indels 22; Gaps 4;

QY 1 HGIFMDRLAS-----KKLCADDECVTYISLASAQEDYNAPDC---RFINVKKQQIY 49

Db 477 HNLLDYNSPKISDFGLAKLCARDQSIVTLTAAGTGMGYIAPELYSRNFGEISYKSDVY 536

QY 50 VYSKLVKENGAGFEWAGVYGDQDEMGVVGYPFR-----NLVKEQRYVQEAATKEVPTT 103

Db 537 SFGMLVL-----EMVSGRRNSDPVESQNVVYFPPEWYIEQVNSGODLALGREMTQEEKET 591

Search completed: September 23, 2002, 09:41:26

Job time: 246 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2002, 09:36:50 ; Search time 30.16 Seconds
(without alignments)
405.110 Million cell updates/sec

Title: US-10-019-455A-24

Perfect score: 590

Sequence: 1 HGIFMDRLASKKLCADCEV.....RVQEATKEVPTTIDIDFCE 110

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

al number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_032802.*
- 1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
 - 2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
 - 3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
 - 4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
 - 5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
 - 6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
 - 7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
 - 8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
 - 9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
 - 10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
 - 11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
 - 12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
 - 13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
 - 14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
 - 15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
 - 16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
 - 17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
 - 18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
 - 19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
 - 20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
 - 21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
 - 22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	590	100.0	110	22	Human MLP protein
2	590	100.0	128	22	Human growth regul
3	590	100.0	128	22	Human MLP protein
4	561	95.1	105	22	Human growth regul
5	547	92.7	110	22	Mouse MLP protein
6	547	92.7	128	22	Mouse MLP protein
7	541	91.7	110	22	Rat MLP protein se
8	541	91.7	128	22	Rat MLP protein se
9	417	70.7	87	22	Rat MLP protein se
10	253.5	43.0	131	16	Melanoma inhibiti
11	253.5	43.0	131	22	Human MIA protein

12	251.5	42.6	137	22	AAG65615	Recombinant human
13	231.5	39.2	130	16	AAR69812	Melanoma inhibiti
14	226.5	38.4	138	22	AAM25834	Human protein sequ
15	225.5	38.2	410	21	AA170210	Human TANGO 130 pr
16	217	36.8	499	22	AAU29319	Human PRO polypept
17	215	36.4	303	22	AAU29304	Human PRO polypept
18	215	36.4	303	22	AAB87608	Human PRO19670..H
19	215	36.4	714	21	AA170209	Murine TANGO 130 p
20	185.5	31.4	1369	20	AA124788	Human secreted pro
21	185.5	31.4	1369	22	AAU39009	Human secreted pro
22	179.5	30.4	114	21	AAG03662	Human secreted pro
23	179.5	30.4	192	22	AAU14154	Human novel protei
24	167	28.3	90	22	AAU14390	Human novel protei
25	157.5	26.7	98	21	AA184900	A human proliferat
26	100	16.9	21	22	AAB69128	Mouse MLP peptide
27	88.5	15.0	318	19	AAB69429	Human secreted pro
28	86.5	14.7	287	20	AA122236	Human KDR signal t
29	86.5	14.7	797	20	AA127125	Amino acid sequenc
30	86.5	14.7	845	21	AA149419	PKA substrate, Vav
31	86.5	14.7	847	20	AA122237	Human KDR signal t
32	85.5	14.5	844	13	AAR25671	Mouse vav proto on
33	82.5	14.0	498	22	AAB95643	Human protein sequ
34	79.5	13.5	186	22	AB10242	Human cDNA seq ID
35	78.5	13.3	1589	21	AA168822	Amino acid sequenc
36	72	12.2	462	17	AAW05395	Human SH3P17 prote
37	72	12.2	641	20	AA132158	Human SH3D1A prote
38	72	12.2	1144	20	AA132154	Human SH3D1A prote
39	71	12.0	2091	21	AA12000	Rat p103 protein.
40	70.5	11.9	663	19	AAW49872	Thermotoga maritim
41	70.5	11.9	680	18	AAW34564	Thermotoga maritim
42	70.5	11.9	680	19	AAW49868	Thermotoga maritim
43	70.5	11.9	680	19	AAW35005	Thermotoga maritim
44	70	11.9	280	19	AAW77273	Protease PD498 mut
45	70	11.9	1683	21	AA171160	Rat phosphodiester

ALIGNMENTS

RESULT 1

AAAB69126

ID AAB69126 standard; Protein; 110 AA.

AC AAB69126;

XX 23-APR-2001 (first entry)

DT Human MLP protein sequence SEQ ID NO:24.

DE MLP: MIA; melanoma inhibitory activity; cancer; bone disease.

KW joint disease; pathologic angiogenesis; diagnosis; antiinflammatory;

KW cardiant; gene therapy; secretory cell function regulator; promoter;

KW inhibitor.

OS Homo sapiens.

XX WO200102564-A1.

PN 11-JAN-2001.

XX 29-JUN-2000; 2000WO-JP04278.

XX 30-JUN-1999; 95JP-0186718.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Itoh Y, Nishi K, Ogi K, Ohkubo S, Mogi S, Noguchi Y, Yoshimura K;

PI Tanaka H;

XX WPI; 2001-159271/16.

DR N-PSDB; AAF59079.

XX Safe, low-toxicity secretory cell function-regulatory protein and

PT encoded DNA, applicable as drugs, in diagnosis and development of
PT promoters and inhibitors for preventing or treating e.g. bone and joint
PT diseases -

PS Claim 1; Page 97-98; 111pp; Japanese.

XX The present invention describes novel MLP proteins and their encoding
CC DNAs. The MLP proteins and DNAs have antiinflammatory and cardiant
CC activities, and can be used in gene therapy and as secretory cell
CC function regulators. The MLP proteins and DNAs can be used in drugs, in
CC the diagnosis and development of promoters and inhibitors for preventing
CC or treating bone and joint diseases as well as pathologic angiogenesis.
CC AAF59063 to AAF59099 and AAB69122 to AAB69132 represent sequences used
CC in the exemplification of the present invention.

XX Sequence 110 AA;

Query Match 100.0%; Score 590; DB 22; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.1e-68;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGIFMDRLASKKLCADDECVTYISLASAQEDYNAPDCRFINVKKGQIYVYSKLVKENG 60

Db 1 hgifmdrlaskklcaddecvtyslasaqedynapdcrfinvkkgqiyvysklvkenga 60

Qy 61 GEFWAGSVYGDQDEMGVVGYPFRLVKEQRYQATKEVPTTDDFFCE 110

Db 61 gefwagsvygdqdemgvvgypfprnlvkeqrvyqeatkevpttdidffce 110

RESULT 2

AAB82671
ID AAB82671 standard; Protein; 128 AA.

AC AAB82671;

DT 02-OCT-2001 (first entry)

DE Human growth regulatory-like polypeptide.

KW Growth regulatory-like polypeptide; human; cartilage; melanoma;
KW neuroectodermal tumour; glioma; cancer; therapy; diagnosis.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Peptide 1..23

FT Protein /label= Signal_peptide

FT /label= Mature_protein

FT /note= "separately claimed in Claim 10"

PN WO200155332-A2.

PD 02-AUG-2001.

PF 25-JAN-2001; 2001WO-US02455.

PR 25-JAN-2000; 2000US-0491404.

PR 02-MAY-2000; 2000US-0563786.

PA (HYSE-) HYSEQ INC.

PI Mize NK, Boyle BJ, Ford JE, Arterburn MC, Tang YT, Liu C;
PI Drmanac RT;

DR WPI; 2001-483233/52.

DR N-PSDB; AAB26343;

XX Isolated human growth regulatory-like polypeptide useful for treating
PT e.g. Alzheimer's disease, cancer, autoimmune disorders,
PT hyperproliferative disorders, coagulation disorders, and nervous system

PT disorders -

PS Claim 10; Page 116-117; 119pp; English.

XX The present sequence is that of a novel human growth regulatory-like
CC polypeptide (GRLP). The amino acid sequence is predicted from a
CC novel assembled cDNA (see AAB26343) based on Hyseq clone number
CC 16372272. The protein has a mol.wt. of 14 kDa unglycosylated. GRLP
CC belongs to the same protein family as growth regulatory proteins,
CC growth factors, human melanoma derived growth regulatory protein
CC precursor (64% similarity and 45% identity over 111 amino acids)
CC or melanoma inhibitory activity, cattle cartilage-derived
CC retinoic acid sensitive protein (CD-RAP, 44% identity and 64%
CC similarity over 126 amino acids) and other retinoic acid-sensitive
CC proteins. GRLP polypeptides and polynucleotides of the invention
CC can be used in the prophylaxis, treatment (including gene therapy)
CC and diagnosis of disorders and diseases caused by, or involving,
CC cartilage development and maintenance, inhibition of melanoma cell
CC growth and tumours, including neuroectodermal tumours such as
CC gliomas. The polypeptides, which include the GRLP mature protein,
CC may also have nutritional uses, cytokine and cell proliferation,
CC or differentiation activity, stem cell growth factor activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC immunosuppressive or immunostimulant activity, activin/inhibin
CC activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, use in cancer diagnosis and therapy,
CC drug screening, receptor/ligand activity, antiinflammatory
CC activity, and treatment of leukaemia, nervous system disorders,
CC arthritis and inflammation.

XX Sequence 128 AA;

Query Match 100.0%; Score 590; DB 22; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.4e-68;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGIFMDRLASKKLCADDECVTYISLASAQEDYNAPDCRFINVKKGQIYVYSKLVKENG 60

Db 19 hgifmdrlaskklcaddecvtyslasaqedynapdcrfinvkkgqiyvysklvkenga 78

Qy 61 GEFWAGSVYGDQDEMGVVGYPFRLVKEQRYQATKEVPTTDDFFCE 110

Db 79 gefwagsvygdqdemgvvgypfprnlvkeqrvyqeatkevpttdidffce 128

RESULT 3

AAB69123
ID AAB69123 standard; Protein; 128 AA.

AC AAB69123;

DT 23-APR-2001 (first entry)

DE Human MLP protein sequence SEQ ID NO:6.

KW MLP; MIA; melanoma inhibitory activity; cancer; bone disease;
KW joint disease; pathologic angiogenesis; diagnosis; antiinflammatory;
KW cardiant; gene therapy; secretory cell function regulator; promoter;
KW inhibitor.

OS Homo sapiens.

PN WO200102564-A1.

PD 11-JAN-2001.

PF 29-JUN-2000; 2000WO-JP04278.

PR 30-JUN-1999; 99JP-0186718.

PA (TAKE) TAKEDA CHEM IND LTD.

XX

PI Itoh Y, Nishi K, Ogi K, Ohkubo S, Mogi S, Noguchi Y, Yoshimura K;
 XX Tanaka H;
 DR WPI: 2001-159271/16.
 DR N-PSDB; AAF59065.
 XX
 PT Safe, low-toxicity secretory cell function-regulatory protein and
 PT encoded DNA, applicable as drugs, in diagnosis and development of
 PT promoters and inhibitors for preventing or treating e.g. bone and joint
 PT diseases
 XX
 XX Claim 2; Page 91-92; 111pp; Japanese.
 XX
 CC The present invention describes novel MLP proteins and their encoding
 CC DNAs. The MLP proteins and DNAs have antiinflammatory and cardiant
 CC activities, and can be used in gene therapy and as secretory cell
 CC function regulators. The MLP proteins and DNAs can be used in drugs, in
 CC the diagnosis and development of promoters and inhibitors for preventing
 CC or treating bone and joint diseases as well as pathologic angiogenesis.
 CC AAF59063 to AAF59099 and AAB69122 to AAB69132 represent sequences used
 CC in the exemplification of the present invention.
 XX
 SQ Sequence 128 AA;
 Query Match 100.0%; Score 590; DB 22; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1.4e-68;
 Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HGIFMDRLASKKLCADDECYVTISLASAOEDYNAPDCRFINVKKGQOIYVSKLVKNGA 60
 DB 19 hgifmdrlaskklcaddecyvtislasaqedynapdcrfinvkkgqgiyvsklvknga 78
 QY 61 GEFWAGSVYGDQDEMVGVPFRNLVKEQRYQVQATKEVPTTIDIDFCE 110
 DB 79 gefwagsvygdqdemvgvypfrnlvkeqrvyqeatkevpttdidfice 128
 RESULT 4
 ID AAB82672
 AC AAB82672 standard; Protein; 105 AA.
 XX
 XX AAB82672;
 XX
 DT 02-OCT-2001 (first entry)
 XX
 DE Human growth regulatory-like polypeptide (mature protein).
 Growth regulatory-like polypeptide; human; cartilage; melanoma;
 neuroectodermal tumour; glioma; cancer; therapy; diagnosis.
 XX
 OS Homo sapiens.
 XX
 PN WO200155332-A2.
 XX
 PD 02-AUG-2001.
 XX
 XX 25-JAN-2001; 2001WO-US02455.
 XX
 XX 25-JAN-2000; 2000US-0491404.
 PR
 PR 02-MAY-2000; 2000US-0563786.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Mize NK, Boyle BJ, Ford JE, Arterburn MC, Tang YT, Liu C;
 PI Drmanac RT;
 PI
 XX WPI: 2001-483233/52.
 DR N-PSDB; AAB26343.
 DR
 XX Isolated human growth regulatory-like polypeptide useful for treating
 PT e.g. Alzheimer's disease, cancer, autoimmune disorders,
 PT hyperproliferative disorders, coagulation disorders, and nervous system

PT disorders -
 XX
 PS Claim 10; Page 117; 119pp; English.
 XX
 CC The present sequence is that of a novel human growth regulatory-like
 CC polypeptide (GRLP) mature protein. The sequence is predicted from
 CC a novel assembled cDNA (see AAB26343) based on Hyseq clone number
 CC 16372272. The protein has a mol.wt. of 14 kDa unglycosylated. GRLP
 CC belongs to the same protein family as growth regulatory proteins,
 CC growth factors, human melanoma derived growth regulatory protein,
 CC precursor (64% similarity and 45% identity over 111 amino acids)
 CC or melanoma inhibitory activity, cattle cartilage-derived
 CC retinoic acid sensitive protein (CD-RAP, 44% identity and 64%
 CC similarity over 126 amino acids) and other retinoic acid-sensitive
 CC proteins. GRLP polypeptides and polynucleotides of the invention
 CC can be used in the prophylaxis, treatment (including gene therapy)
 CC and diagnosis of disorders and diseases caused by, or involving,
 CC cartilage development and maintenance, inhibition of melanoma cell
 CC growth and tumours, including neuroectodermal tumours such as
 CC gliomas. The polypeptides, which include the GRLP mature protein,
 CC may also have nutritional uses, cytokine and cell proliferation
 CC or differentiation activity, stem cell growth factor activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC immunosuppressive or immunostimulant activity, activin/inhibin
 CC activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, use in cancer diagnosis and therapy,
 CC drug screening, receptor/ligand activity, antiinflammatory
 CC activity, and treatment of leukaemia, nervous system disorders,
 CC arthritis and inflammation.
 XX
 SQ Sequence 105 AA;
 Query Match 95.1%; Score 561; DB 22; Length 105;
 Best Local Similarity 100.0%; Pred. No. 6.2e-65;
 Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 DRLASKKLCADDECYVTISLASAOEDYNAPDCRFINVKKGQOIYVSKLVKENGAEFFWA 65
 DB 1 drlaskklcaddecyvtislasaqedynapdcrfinvkkgqgiyvsklvkengaeffwa 60
 QY 66 GSVYGDQDEMVGVPFRNLVKEQRYQVQATKEVPTTIDIDFCE 110
 DB 61 gsvygdqdemvgvypfrnlvkeqrvyqeatkevpttdidfice 105
 RESULT 5
 AAB69127
 ID AAB69127 standard; Protein; 110 AA.
 XX
 XX AAB69127;
 XX
 DT 23-APR-2001 (first entry)
 XX
 DE Mouse MLP protein sequence SEQ ID NO:26.
 XX
 KW MLP; MTA; melanoma inhibitory activity; cancer; bone disease;
 KW joint disease; pathologic angiogenesis; diagnosis; antiinflammatory;
 KW cardiant; gene therapy; secretory cell function regulator; promoter;
 KW inhibitor.
 XX
 OS Mus musculus.
 XX
 XX WO200102564-A1.
 PN
 XX 11-JAN-2001.
 PD
 XX 29-JUN-2000; 2000WO-JP04278.
 PF
 XX 30-JUN-1999; 99JP-0186718.
 PR
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA
 XX

PI Itoh Y, Nishi K, Ogi K, Ohkubo S, Mogi S, Noguchi Y, Yoshimura K;
PI Tanaka H;
XX WPI: 2001-159271/16.
DR N-PSDB; AAF59080.
XX
PT Safe, low-toxicity secretory cell function-regulatory protein and
PT encoded DNA, applicable as drugs, in diagnosis and development of
PT promoters and inhibitors for preventing or treating e.g. bone and joint
PT diseases -
XX
XX Claim 3; Page 98-99; 111pp; Japanese.
XX
CC The present invention describes novel MLP proteins and their encoding
CC DNAs. The MLP proteins and DNAs have antiinflammatory and cardiant
CC activities, and can be used in gene therapy and as secretory cell
CC function regulators. The MLP proteins and DNAs can be used in drugs, in
CC the diagnosis and development of promoters and inhibitors for preventing
CC or treating bone and joint diseases as well as pathologic angiogenesis.
CC AAF59063 to AAF59099 and AAB69122 to AAB69132 represent sequences used
CC in the exemplification of the present invention.
XX
XX Sequence 110 AA;
XX
XX Query Match 92.7%; Score 547; DB 22; Length 110;
XX Best Local Similarity 90.0%; Pred. No. 4.4e-63;
XX Matches 99; Conservative 7; Mismatches 4; Indels 0; Gaps 0;
XX
QY 1 HGTFMDRLASKKLCADDECVYTISLASAQEDYNAPDCRFNVKKGQIYVYKLVKNGA 60
DB 1 hgvmfmdklskklcadeecvvtislaragedynapdcridvkgqgiyvyksklvtenga 60
QY 61 GEFWAGSVYGDGDEMGVYGFPRNLVKEQRYVQATKEVPTTDIDFFCE 110
DB 61 gefwagsvygdhqdengivgyfprnlvkeqrvyqeatkeipttdidffce 110
XX
RESULT 6
AAB69125
ID AAB69125 standard; Protein; 128 AA.
XX
XX AAB69125;
AC
XX
DT 23-APR-2001 (first entry)
XX
XX Mouse MLP protein sequence SEQ ID NO:12.
DE
XX
XX MLP; MTA; melanoma inhibitory activity; cancer; bone disease;
XX joint disease; pathologic angiogenesis; diagnosis; antiinflammatory;
XX cardiant; gene therapy; secretory cell function regulator; promoter;
XX inhibitor.
XX
XX Mus musculus.
OS
XX
XX WO200102564-A1.
PN
XX
XX 11-JAN-2001.
PD
XX
XX 29-JUN-2000; 2000WO-JP04278.
PF
XX
XX 30-JUN-1999; 99JP-0186718.
PR
XX
XX (TAKE) TAKEDA CHEM IND LTD.
PA
XX
XX Itoh Y, Nishi K, Ogi K, Ohkubo S, Mogi S, Noguchi Y, Yoshimura K;
PI Tanaka H;
XX
XX WPI: 2001-159271/16.
DR
XX
XX N-PSDB; AAF59068.
DR
XX
XX Safe, low-toxicity secretory cell function-regulatory protein and
XX encoded DNA, applicable as drugs, in diagnosis and development of
XX promoters and inhibitors for preventing or treating e.g. bone and joint
XX diseases -
XX
XX Claim 5; Page 107; 111pp; Japanese.
XX
CC The present invention describes novel MLP proteins and their encoding
CC DNAs. The MLP proteins and DNAs have antiinflammatory and cardiant
CC activities, and can be used in gene therapy and as secretory cell

PT promoters and inhibitors for preventing or treating e.g. bone and joint
PT diseases -
XX
XX Claim 4; Page 93-94; 111pp; Japanese.
XX
CC The present invention describes novel MLP proteins and their encoding
CC DNAs. The MLP proteins and DNAs have antiinflammatory and cardiant
CC activities, and can be used in gene therapy and as secretory cell
CC function regulators. The MLP proteins and DNAs can be used in drugs, in
CC the diagnosis and development of promoters and inhibitors for preventing
CC or treating bone and joint diseases as well as pathologic angiogenesis.
CC AAF59063 to AAF59099 and AAB69122 to AAB69132 represent sequences used
CC in the exemplification of the present invention.
XX
XX Sequence 128 AA;
XX
XX Query Match 92.7%; Score 547; DB 22; Length 128;
XX Best Local Similarity 90.0%; Pred. No. 5.4e-63;
XX Matches 99; Conservative 7; Mismatches 4; Indels 0; Gaps 0;
XX
QY 1 HGTFMDRLASKKLCADDECVYTISLASAQEDYNAPDCRFNVKKGQIYVYKLVKNGA 60
DB 19 hgvmfmdklskklcadeecvvtislaragedynapdcridvkgqgiyvyksklvtenga 78
QY 61 GEFWAGSVYGDGDEMGVYGFPRNLVKEQRYVQATKEVPTTDIDFFCE 110
DB 79 gefwagsvygdhqdengivgyfprnlvkeqrvyqeatkeipttdidffce 128
XX
RESULT 7
AAB69131
ID AAB69131 standard; Protein; 110 AA.
XX
XX AAB69131;
AC
XX
DT 23-APR-2001 (first entry)
XX
XX Rat MLP protein sequence SEQ ID NO:49.
DE
XX
XX MLP; MTA; melanoma inhibitory activity; cancer; bone disease;
XX joint disease; pathologic angiogenesis; diagnosis; antiinflammatory;
XX cardiant; gene therapy; secretory cell function regulator; promoter;
XX inhibitor.
XX
XX Rattus sp.
OS
XX
XX WO200102564-A1.
PN
XX
XX 11-JAN-2001.
PD
XX
XX 29-JUN-2000; 2000WO-JP04278.
PF
XX
XX 30-JUN-1999; 99JP-0186718.
PR
XX
XX (TAKE) TAKEDA CHEM IND LTD.
PA
XX
XX Itoh Y, Nishi K, Ogi K, Ohkubo S, Mogi S, Noguchi Y, Yoshimura K;
PI Tanaka H;
XX
XX WPI: 2001-159271/16.
DR
XX
XX N-PSDB; AAF59099.
DR
XX
XX Safe, low-toxicity secretory cell function-regulatory protein and
XX encoded DNA, applicable as drugs, in diagnosis and development of
XX promoters and inhibitors for preventing or treating e.g. bone and joint
XX diseases -
XX
XX Claim 5; Page 107; 111pp; Japanese.
XX
CC The present invention describes novel MLP proteins and their encoding
CC DNAs. The MLP proteins and DNAs have antiinflammatory and cardiant
CC activities, and can be used in gene therapy and as secretory cell

QY 66 GSVYGDQDEMGVGVYFPRNLVKEQRY 92
 Db 61 gsvygdqdemgvgvypsnlvregrv 87

RESULT 10

AAR69811
 ID AAR69811 standard; Protein; 131 AA.

XX AAR69811;
 AC AAR69811;

DT 26-OCT-1995 (first entry)
 XX

DE Melanoma inhibiting protein (human).
 XX

XX Melanoma; inhibition; cancer; melanoma; glioblastoma; neuroblastoma;
 KW small cell lung cancer; neuroectodermal tumours; immunosuppressant;
 KW phytohaemagglutinin; lymphocyte; interleukin 2; IL-2; detection;
 KW probe; fusion protein.
 XX

XX Homo sapiens.
 S

XX WO9503328-A.
 N

XX 02-FEB-1995.
 PD

XX 19-JUL-1994; 94WO-EP02369.
 PF

XX 20-JUL-1993; 93DE-4324247.
 PR

XX (BOEF) BOEHRINGER MANNHEIM GMBH.
 PA

XX Bogdahn U, Buettner R, Kaluza B;
 PI

XX WPI: 1995-075191/10.
 DR

XX N-PSDB; AAQ84050, AAQ84051.
 DR

XX New melanoma inhibiting protein and related nucleic acid -
 PT vectors, transformed cells, antibodies etc., useful for treating
 PT tumours and as immunosuppressant e.g. by gene therapy
 XX

PS Claim 1; Page 54; 85pp; German.
 PS

XX This protein has melanoma-inhibiting activity and can be used to
 CC treat cancer (melanoma, glioblastoma, neuroblastoma, small cell
 CC lung cancer, neuroectodermal tumours) or as an immunosuppressant
 CC (it inhibits IL-2 or phytohaemagglutinin induced proliferation of
 CC peripheral blood lymphocytes). Antibodies raised against the
 CC protein can be used to detect cell producing the protein and also
 CC for protein purification. Probes derived from DNA encoding the
 CC protein (AAQ84050, AAQ84051) can be used to detect sequences encoding
 CC the protein or related proteins. The protein may be expressed as
 CC a fusion protein (conjugated with dihydrofolate reductase (DHFR)).
 XX

SQ Sequence 131 AA;

Query Match 43.0%; Score 253.5; DB 16; Length 131;
 Best Local Similarity 45.4%; Pred. No. 8.7e-25;
 Matches 49; Conservative 21; Mismatches 33; Indels 5; Gaps 3;

QY 5 MDRLASKKLCADDECVTYISLASAEDYNAPDCRFINVKKGQOIYVYSKLVKENGAGE-F 63
 Db 27 mpkladrklcadqecshpismavalqdympdcrcfithrgqvvyvfskl---kgrgrlf 83

QY 64 WAGSVYGDQDEMGV-VGYFPRNLVKEQRYOEATKEVPTTIDDFCE 110
 Db 84 wggsvggdygdlarigvfpssivredqclpgkvdvktkwdlyc 131

RESULT 11

AAG65614
 ID AAG65614 standard; Protein; 131 AA.

XX AAG65614;
 AC

DT 07-JAN-2002 (first entry)
 DT

DE Human MIA protein sequence.
 DE

XX MIA; melanoma inhibiting activity protein; antiinflammatory; human;
 KW antiarthritic; antirheumatic; antithyroid; osteopathic; nephrotropic;
 KW immunosuppressive; ophthalmological; dermatological; antidiabetic;
 KW neuroprotective; immune tolerance; T-cell tolerance.
 XX

OS Homo sapiens.
 OS

XX WO200170253-A1.
 PN

XX 27-SEP-2001.
 PD

XX 15-MAR-2001; 2001WO-EP02991.
 PF

XX 23-MAR-2000; 2000EP-0201063.
 PR

XX (ALKU) AKZO NOBEL NV.
 XX

PA Nelissen RLH, Verheijden GFM;
 PI

XX WPI: 2001-611446/70.
 DR

XX N-PSDB; AAH47783.
 DR

XX Use of melanoma inhibiting activity protein or its derivatives as
 PT immune modulatory agents for the treatment of inflammatory diseases,
 PT specifically rheumatoid arthritis -
 PT

PS Example 4; Page 34; 41pp; English.
 PS

XX The invention relates to the use of melanoma inhibiting activity (MIA)
 CC protein and/or its fragments that have anti-inflammatory effects and
 CC induce systemic immune tolerance or specific T-cell tolerance to MIA
 CC antigen, for manufacturing a preparation against inflammatory diseases
 CC and for induction of systemic immune tolerance or specific T-cell
 CC tolerance in patients suffering from or susceptible to inflammatory
 CC diseases. A fragment of MIA is useful as a therapeutic substance and is
 CC useful for manufacture of pharmaceutical preparations against
 CC inflammatory diseases such as an immune-cell mediated cartilage
 CC destruction disease, specifically rheumatoid arthritis, autoimmune
 CC diseases like Graves' disease, juvenile arthritis, primary
 CC glomerulonephritis, polyarthritis, osteoarthritis, Sjogren's syndrome,
 CC myasthenia gravis, Addison's disease, primary biliary sclerosis,
 CC uveitis, systemic lupus erythematosus, inflammatory bowel disease,
 CC multiple sclerosis or diabetes. The MIA polypeptides have specific
 CC effect on the autoreactive T-cells thus leaving the other components of
 CC the immune system intact as compared to the non-specific suppressive
 CC effect of immunosuppressive drugs. The present sequence represents a
 CC human MIA protein.
 XX

SQ Sequence 131 AA;

Query Match 43.0%; Score 253.5; DB 22; Length 131;
 Best Local Similarity 45.4%; Pred. No. 8.7e-25;
 Matches 49; Conservative 21; Mismatches 33; Indels 5; Gaps 3;

QY 5 MDRLASKKLCADDECVTYISLASAEDYNAPDCRFINVKKGQOIYVYSKLVKENGAGE-F 63
 Db 27 mpkladrklcadqecshpismavalqdympdcrcfithrgqvvyvfskl---kgrgrlf 83

QY 64 WAGSVYGDQDEMGV-VGYFPRNLVKEQRYOEATKEVPTTIDDFCE 110
 Db 84 wggsvggdygdlarigvfpssivredqclpgkvdvktkwdlyc 131

RESULT 12

AAG65615

ID AAG5615 standard; Protein; 137 AA.
 XX AAG5615;
 AC
 XX
 DT 07-JAN-2002 (first entry)
 DE Recombinant human MIA(his7) protein.
 XX
 XX MIA; melanoma inhibiting activity protein; antiinflammatory; human;
 KW antarthritic; antirheumatic; antithyroid; osteopathic; nephrotropic;
 KW immunosuppressive; ophthalmological; dermatological; antidiabetic;
 KW neuroprotective; immune tolerance; T-cell.
 XX
 OS Homo sapiens.
 XX
 XX WO200170253-A1.
 PN
 XX
 PD 27-SEP-2001.
 XX
 XX 15-MAR-2001; 2001WO-EP02991.
 PR
 XX 23-MAR-2000; 2000EP-0201063.
 PA (ALKU) AKZO NOBEL NV.
 XX
 XX Nelissen RLH, Verheijden GPM;
 PI
 XX WPI; 2001-611446/70.
 DR N-PSDB; AAH47783.
 XX
 XX Use of melanoma inhibiting activity protein or its derivatives as
 PT immune modulatory agents for the treatment of inflammatory diseases,
 PT specifically rheumatoid arthritis -
 XX
 XX Example 1; Page 35; 41pp; English.
 XX
 CC The invention relates to the use of melanoma inhibiting activity (MIA)
 CC protein and/or its fragments that have anti-inflammatory effects and
 CC induce systemic immune tolerance or specific T-cell tolerance to MIA
 CC antigen, for manufacturing a preparation against inflammatory diseases
 CC and for induction of systemic immune tolerance or specific T-cell
 CC tolerance in patients suffering from or susceptible to inflammatory
 CC diseases. A fragment of MIA is useful as a therapeutic substance and is
 CC useful for manufacture of pharmaceutical preparations against
 CC inflammatory diseases such as an immune-cell mediated cartilage
 CC destruction disease, specifically rheumatoid arthritis, autoimmune
 CC diseases like Graves' disease, juvenile arthritis, primary
 CC glomerulonephritis, polyarthritis, osteoarthritis, Sjogren's syndrome,
 CC myasthenia gravis, Addison's disease, primary biliary sclerosis,
 CC uveitis, systemic lupus erythematosus, inflammatory bowel disease,
 CC multiple sclerosis or diabetes. The MIA polypeptides have specific
 CC effect on the autoreactive T-cells thus leaving the other components of
 CC the immune system intact as compared to the non-specific suppressive
 CC effect of immunosuppressive drugs. The present sequence represents a
 CC cDNA encoding a human recombinant MIA(his7) protein.
 XX
 XX Sequence 137 AA;
 SQ
 Query Match 42.6%; Score 251.5; DB 22; Length 137;
 Best Local Similarity 45.8%; Pred. No. 1.7e-24;
 Matches 49; Conservative 20; Mismatches 33; Indels 5; Gaps 3;
 QY 5 MDRLASKKLCADCECVYTISLASAQEDYNAPDCRFNNVKKGOIYVYSLVKENGAGE-F 63
 Db 27 mpkladkicadecqshpismavaldymapdcrfnhrvgqvvyfskl---kggrlf 83
 QY 64 WAGSVYGDGDEMGV-VGYFPPRLNFKQRYQVQATKEVPTDIDFFC 109
 Db 84 wgsvvgdygdlarlgypfssivredgtlkpgkdvtkdkwdfc 130
 RESULT 13
 Human protein sequence SEQ ID NO:1349.

AAR69812
 ID AAR69812 standard; Protein; 130 AA.
 XX
 AC AAR69812;
 XX
 DT 26-OCT-1995 (first entry)
 DE Melanoma inhibiting protein (murine).
 XX
 XX Melanoma; inhibition; cancer; melanoma; glioblastoma; neuroblastoma;
 KW small cell lung cancer; neuroectodermal tumours; immunosuppressant;
 KW phytohaemagglutinin; lymphocyte; interleukin 2; IL-2; detection;
 KW probe; fusion protein.
 XX
 OS Mus musculus.
 XX
 XX WO9503328-A.
 PN
 XX
 PD 02-FEB-1995.
 XX
 XX 19-JUL-1994; 94WO-EP02369.
 PF
 XX 20-JUL-1993; 93DE-4324247.
 PR
 XX (BOEF) BOEHRINGER MANNHEIM GMBH.
 PA
 XX Bogdahn U, Buettner R, Kaluza B;
 PI
 XX WPI; 1995-075191/10.
 DR N-PSDB; AAQ45052.
 XX
 XX New melanoma inhibiting protein and related nucleic acid -
 PT vectors, transformed cells, antibodies etc., useful for treating
 PT tumours and as immunosuppressant e.g. by gene therapy
 XX
 XX Claim 5; Page 60; 85pp; German.
 PS
 CC This protein has melanoma-inhibiting activity and can be used to
 CC treat cancer (melanoma, glioblastoma, neuroblastoma, small cell
 CC lung cancer, neuroectodermal tumours) or as an immunosuppressant
 CC (it inhibits IL-2 or phytohaemagglutinin induced proliferation of
 CC peripheral blood lymphocytes). Antibodies raised against the
 CC protein can be used to detect cell producing the protein and also
 CC for protein purification. Probes derived from DNA encoding the
 CC protein (AAQ84052) can be used to detect sequences encoding the
 CC protein or related proteins. The protein may be expressed as
 CC a fusion protein (conjugated with dihydrofolate reductase (DHFR)).
 XX
 XX Sequence 130 AA;
 SQ
 Query Match 39.2%; Score 231.5; DB 16; Length 130;
 Best Local Similarity 43.5%; Pred. No. 6.3e-22;
 Matches 47; Conservative 21; Mismatches 35; Indels 5; Gaps 3;
 QY 5 MDRLASKKLCADCECVYTISLASAQEDYNAPDCRFNNVKKGOIYVYSLVKENGAGE-F 63
 Db 26 mpkladkicadecqshpismavaldymapdcrfnhrvgqvvyfskl---kggrlf 82
 QY 64 WAGSVYGDGDEMGV-VGYFPPRLNFKQRYQVQATKEVPTDIDFFC 110
 Db 83 wgsvvgdygdlarlgypfssivredgtlnskpgkldmktcdqwdqfyc 130
 RESULT 14
 AAM25834
 ID AAM25834 standard; Protein; 138 AA.
 XX
 AC AAM25834;
 XX
 DT 16-OCT-2001 (first entry)
 DE Human protein sequence SEQ ID NO:1349.

CC factor function and embryonic cell growth/morphogenesis.

XX	Sequence	410 AA;
SQ		

Query Match	38.2%;	Score 225.5;	DB 21;	Length 410;
Best Local Similarity	44.7%;	Pred. No. 1.9e-20;		
Matches 46;	Conservative 14;	Mismatches 34;	Indels 9;	Gaps 3;

Qy 7 RLASKKLCADDECVTISLSAQEDYNAPDCRFINVKKGQIYVYSLVKENGAGEFWAG 66

```

Db      | : ||||| | : ||: ||||: ||: ||||: | | ||
31 rfsehkicaddecsmlyrgaeledftopdcrfvntkqdpvyvykylar--gwpewag 88

```

Qy 67 SVYGDQDEMVGVFPRNLVKEQRVQEATKEVPTTIDFFC 109

```

Db      || | : ||||:|: | : ||| : || |
89 sv---gr----tfqyfpkdlqvvheytkeelqvptdetc 124

```

Search completed: September 23, 2002, 09:37:53
 Elapsed time: 63 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2002, 09:36:50 ; Search time 12.89 Seconds
(without alignments)
208.442 Million cell updates/sec

Title: US-10-019-455A-24
Perfect score: 590
Sequence: 1 HGFMRLASKKLCADCECV.....RVQEQATKEVPTTIDIFCEE 110

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Cal number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
 - 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
 - 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
 - 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
 - 5: /cgn2_6/ptodata/2/1aa/PTCUTS_COMB.pep.*
 - 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	253.5	43.0	131	1 US-08-578-649-2	Sequence 2, Appli
2	231.5	39.2	130	1 US-08-578-649-5	Sequence 5, Appli
3	85.5	14.5	844	1 US-07-646-537B-2	Sequence 2, Appli
4	81.5	13.8	54	4 US-09-346-510B-21	Sequence 21, Appli
5	78.5	13.3	1589	3 US-09-356-952-4	Sequence 4, Appli
6	72.5	12.3	48	4 US-09-346-510B-24	Sequence 24, Appli
7	72	12.2	462	4 US-08-630-915A-38	Sequence 38, Appli
8	66.5	11.3	62	4 US-08-630-915A-135	Sequence 135, Appli
9	66.5	11.3	642	2 US-08-245-511-48	Sequence 48, Appli
10	66.5	11.3	642	2 US-08-600-993A-48	Sequence 48, Appli
11	66	11.2	248	4 US-08-630-915A-40	Sequence 40, Appli
12	66	11.2	280	1 US-08-434-255-8	Sequence 8, Appli
13	66	11.2	280	1 US-08-459-967-8	Sequence 8, Appli
14	66	11.2	280	1 US-08-460-327-6	Sequence 6, Appli
15	66	11.2	280	1 US-08-459-871-8	Sequence 8, Appli
16	66	11.2	280	4 US-09-024-532-2	Sequence 2, Appli
17	66	11.2	280	4 US-09-104-623A-2	Sequence 2, Appli
18	66	11.2	370	1 US-08-434-255-6	Sequence 6, Appli
19	66	11.2	370	1 US-08-459-967-6	Sequence 6, Appli
20	66	11.2	370	1 US-08-460-327-6	Sequence 6, Appli
21	66	11.2	370	1 US-08-459-871-6	Sequence 6, Appli
22	66	11.2	397	1 US-08-434-255-2	Sequence 2, Appli
23	66	11.2	397	1 US-08-434-255-4	Sequence 4, Appli
24	66	11.2	397	1 US-08-459-967-2	Sequence 2, Appli
25	66	11.2	397	1 US-08-459-967-4	Sequence 4, Appli
26	66	11.2	397	1 US-08-460-327-2	Sequence 2, Appli
27	66	11.2	397	1 US-08-460-327-4	Sequence 4, Appli

28 66 11.2 397 1 US-08-459-871-2 Sequence 2, Appli

29 66 11.2 397 1 US-08-459-871-4 Sequence 4, Appli

30 66 11.2 509 4 US-08-630-915A-194 Sequence 194, App

31 65 11.0 57 4 US-08-630-915A-136 Sequence 136, App

32 64 10.8 57 4 US-08-630-915A-219 Sequence 219, App

33 64 10.8 433 1 US-08-700-359-20 Sequence 20, Appl

34 64 10.8 532 3 US-08-911-321-11 Sequence 11, Appl

35 64 10.8 803 1 US-08-158-232-10 Sequence 10, Appl

36 64 10.8 803 1 US-08-304-626-10 Sequence 10, Appl

37 64 10.8 803 1 US-08-316-301A-12 Sequence 12, Appl

38 64 10.8 803 2 US-08-611-928-10 Sequence 10, Appl

39 64 10.8 803 3 US-09-173-891-10 Sequence 10, Appl

40 64 10.8 803 4 US-09-076-137-12 Sequence 12, Appl

41 64 10.8 803 5 PCT-US92-03624-12 Sequence 12, Appl

42 63.5 10.8 58 4 US-08-630-915A-210 Sequence 210, App

43 63 10.7 57 4 US-08-630-915A-125 Sequence 125, App

44 63 10.7 543 4 US-08-426-509A-14 Sequence 14, Appl

45 63 10.7 543 5 PCT-US95-05008-14 Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-08-578-649-2

: Sequence 2, Application US/08578649

: Patent No. 5770366

: GENERAL INFORMATION:

: APPLICANT: Ulrich Bogdan

: APPLICANT: Reinhard Buttner

: APPLICANT: Brigitte Kaluza

: TITLE OF INVENTION: MELANOMA-INHIBITING PROTEIN

: NUMBER OF SEQUENCES: 24

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Felfe & Lynch

: STREET: 805 Third Avenue

: CITY: New York

: STATE: New York

: COUNTRY: USA

: ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/578,649

FILING DATE: 29-July-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE P 43 24 247.2

FILING DATE: 20-July-1993

ATTORNEY/AGENT INFORMATION:

NAME: Andrew L. Tiajolloff

REGISTRATION NUMBER: 31,575

REFERENCE/DOCKET NUMBER: BOER 1035-PFF/ALT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 888-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 131 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-578-649-2

Query Match 43.0%; Score 253.5; DB 1; Length 131;

Best Local Similarity 45.4%; Pred. No. 1.9e-25;

Matches 49; Conservative 21; Mismatches 33; Indels 5; Gaps 3;

QY 5 MDRLASKKLCADCECVTTISLASAQEDYNAPDCRFNVKKGQIYVYVKLVKENGAGE-F 63

Db 27 MPKLADKLCADQECGHPISMAVALQDYNAPDCRFLTIHRGQVYVFSKL---KGRGLF 83
QY 64 WAGSVYGDGQDEMGV-VGYFPRNLVKEQRYQVQATKEVPTTDDIFPCE 110
Db 84 WGSVQGGYDGLAARLGYFPSSIVREDQTLKPGKVDVYTDKWDVFCQ 131

RESULT 2

US-08-578-649-5
; Sequence 5, Application US/08578649
; Patent No. 5770366
; GENERAL INFORMATION:
; APPLICANT: Ulrich Bogdan
; APPLICANT: Reinhard Buttner
; APPLICANT: Brigitte Kaluza
; TITLE OF INVENTION: MELANOMA-INHIBITING PROTEIN
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578,649
; FILING DATE: 29-July-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P.43 24 247.2
; FILING DATE: 20-July-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Andrew L. Tiajolofof
; REGISTRATION NUMBER: 31,575
; REFERENCE/DOCKET NUMBER: BOER 1035-PFF/ALT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 130 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; 3-08-578-649-5

Query Match 39.2%; Score 231.5; DB 1; Length 130;
Best Local Similarity 43.5%; Pred. No. 1.4e-22;
Matches 47; Conservative 21; Mismatches 35; Indels 5; Gaps 3;

QY 5 MDRLASKKLCADCECVYTISLASAODYNAPDCRFLTNVKKGOIYVSKLVKENGAGE-F 63
Db 26 MPKLADKLCADQECGHPISMAVALQDYNAPDCRFLTIHRGQVYVFSKL---KGRGLF 82
QY 64 WAGSVYGDGQDEMGV-VGYFPRNLVKEQRYQVQATKEVPTTDDIFPCE 110
Db 83 WGSVQGGYDGLAARLGYFPSSIVREDLNSKPGKIDMKTDQWDFVCQ 130

RESULT 3

US-07-646-537B-2
; Sequence 2, Application US/07646537B
; Patent No. 5348864
; GENERAL INFORMATION:
; APPLICANT: Barbacid, Mariano
; TITLE OF INVENTION: Vav Proto-Oncogene Protein

; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bristol-Myers Squibb Company
; STREET: P.O. Box 4000
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 08543-4000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/646,537B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaul, Timothy J.
; REGISTRATION NUMBER: 33,111
; REFERENCE/DOCKET NUMBER: DG10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 921-5901
; TELEFAX: (609) 921-4526
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 844 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-646-537B-2

Query Match 14.5%; Score 85.5; DB 1; Length 844;
Best Local Similarity 31.4%; Pred. No. 0.021; 21; Indels 13; Gaps 3;

Matches 22; Conservative 14; Mismatches 14; Indels 13; Gaps 3;

QY 26 ASAEQDYNAPDCRFLTNVKKGOIYVSKLVKENGAGEFWAGSVYGDGQDEMGVYVGYFPRN 85
Db 786 AKARYDFCARDRSLSLKEGDII---KILNKKGGQGGWRRGEIYGR-----IGWFPSN 834

QY 86 LVKEQRYQVE 95
Db 835 YVEED--YSE 842

RESULT 4

US-09-346-510B-21
; Sequence 21, Application US/09346510B
; Patent No. 6281014
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.

; TITLE OF INVENTION: SH3-Containing Protein, DNA and Uses Thereof
; FILE REFERENCE: D6221CIP
; CURRENT APPLICATION NUMBER: US/09/346,510B
; CURRENT FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 08/871,732
; PRIOR FILING DATE: 1997-06-09
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 21
; LENGTH: 54
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; NAME/KEY: Domain
; OTHER INFORMATION: amino acid sequence of Vav SH3 domain
; US-09-346-510B-21

Query Match 13.8%; Score 81.5; DB 4; Length 54;
Best Local Similarity 32.3%; Pred. No. 0.0015;
Matches 20; Conservative 13; Mismatches 18; Indels 11; Gaps 2;

NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
APPLICANT: Klausner & Jackson
ADDRESSEE: 411 Hackensack Avenue
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,993A
FILING DATE: 1-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,511
FILING DATE: 18-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/116,541
FILING DATE: 01-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-069 US
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 642 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORGANISM: Streptococcus pneumoniae
IMMEDIATE SOURCE:
CLONE: amia
FEATURE:
OTHER INFORMATION: the reference contains a sequence error; the
OTHER INFORMATION: correct sequence shown below is obtained from GENBANK
PUBLICATION INFORMATION:
AUTHORS: Allosing, et al.
JOURNAL: Mol. Microbiol.
VOLUME: 4
PAGES: 633-644
DATE: 1990
US-08-600-993A-48

Query Match 11.3%; Score 66.5; DB 2; Length 642;
Best Local Similarity 22.5%; Pred. No. 4.4;
Matches 27; Conservative 20; Mismatches 44; Indels 29; Gaps 4;
QY 6 DRLASKKLCADCECVTISLASAQEDYNAPDCRFIVKKGQIYVYVSKLVKEN-----58
Db 309 EKVAITKALLNKDFRALFALDRSAQO-----INGKDGAAALAVRNLFVKPDFVSAGEK 364
QY 59 -----GAGEFWAGSVYGDGDMGVGYFFRNLYKEQRYOEATKEVPTTIDIF 107
Db 365 TFGDLVAQLPAVGDWKGYNLADGQD-----GLF--NADKAKAEPRKAKALEADGVQF 417
RESULT 11
US-08-630-915A-40
Sequence 40, Application US/08630915A
Patent No. 6309820

GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
ADDRESSEE: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLKES, Dana M.
APPLICANT: McCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
TITLE OF INVENTION: USING SAME
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,972
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 248 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-630-915A-40
Query Match 11.2%; Score 66; DB 4; Length 248;
Best Local Similarity 30.4%; Pred. No. 1.4;
Matches 24; Conservative 7; Mismatches 24; Indels 24; Gaps 3;
QY 28 AQEDYNAPDCRFIVKKGQIYVYVSKLVKENGAGFEWAGSVYGDGDMGVGYFFRNLY 87
Db 190 AMYDYAANNEDELSFSGQLINYNK-----DDPDWQGEI-----NGVTGLFSPSYV 237
QY 88 KEORVYQEATKEVPTTIDID 106
Db 238 K-----MTTDS 244
RESULT 12
US-08-434-255-8
Sequence 8, Application US/08434255
Patent No. 5621089
GENERAL INFORMATION:
APPLICANT: Sloma, Alan P.
APPLICANT: Outtrup, Helle
APPLICANT: Dammann, Claus
APPLICANT: Aaslyng, Dorrit
TITLE OF INVENTION: ALKALINE PROTEASE
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5621089o No. 5621089th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York

; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Agris Dr., Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 3764.400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-434-255-8

Query Match 11.2%; Score 66; DB 1; Length 280;
Best Local Similarity 29.6%; Pred. No. 1.6;
Matches 21; Conservative 11; Mismatches 39; Indels 0; Gaps 0;

QY 22 TISLASAEDYNAPDCRFNVKKGQOIVYYSKLVKENGAGFEWAGSVYGDGQDEMGVVG 81
|::: : ||| || : | : || |::: : |||
Db 34 TVAVLDSGVYDYNHPDLARKVKGDFIDRNNPMDLNGHGHVAGTVAADTNGIGVAGM 93
QY 82 FPRNLVKEQRV 92
| : ||
Db 94 APDTKILAVRV 104

RESULT 13
US-08-459-967-8
; Sequence 8, Application US/08459967
; Patent No. 5622841
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan P.
; APPLICANT: Outtrup, Helle
; APPLICANT: Dambmann, Claus
; APPLICANT: Aslyng, Dorrit
; TITLE OF INVENTION: ALKALINE PROTEASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 56228410 No. 5622841disk of No. 5622841th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,967
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,255
; FILING DATE: 03-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Agris Dr., Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 3764.400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-967-8

; REFERENCE/DOCKET NUMBER: 3764.400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-967-8

Query Match 11.2%; Score 66; DB 1; Length 280;
Best Local Similarity 29.6%; Pred. No. 1.6;
Matches 21; Conservative 11; Mismatches 39; Indels 0; Gaps 0;

QY 22 TISLASAEDYNAPDCRFNVKKGQOIVYYSKLVKENGAGFEWAGSVYGDGQDEMGVVG 81
|::: : ||| || : | : || |::: : |||
Db 34 TVAVLDSGVYDYNHPDLARKVKGDFIDRNNPMDLNGHGHVAGTVAADTNGIGVAGM 93
QY 82 FPRNLVKEQRV 92
| : ||
Db 94 APDTKILAVRV 104

RESULT 14
US-08-460-327-8
; Sequence 8, Application US/08460327
; Patent No. 5622850
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan P.
; APPLICANT: Outtrup, Helle
; APPLICANT: Dambmann, Claus
; APPLICANT: Aslyng, Dorrit
; TITLE OF INVENTION: ALKALINE PROTEASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 56228500 No. 5622850disk of No. 5622850th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,327
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,255
; FILING DATE: 03-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Agris Dr., Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 3764.400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-460-327-8

Query Match 11.2%; Score 66; DB 1; Length 280;

